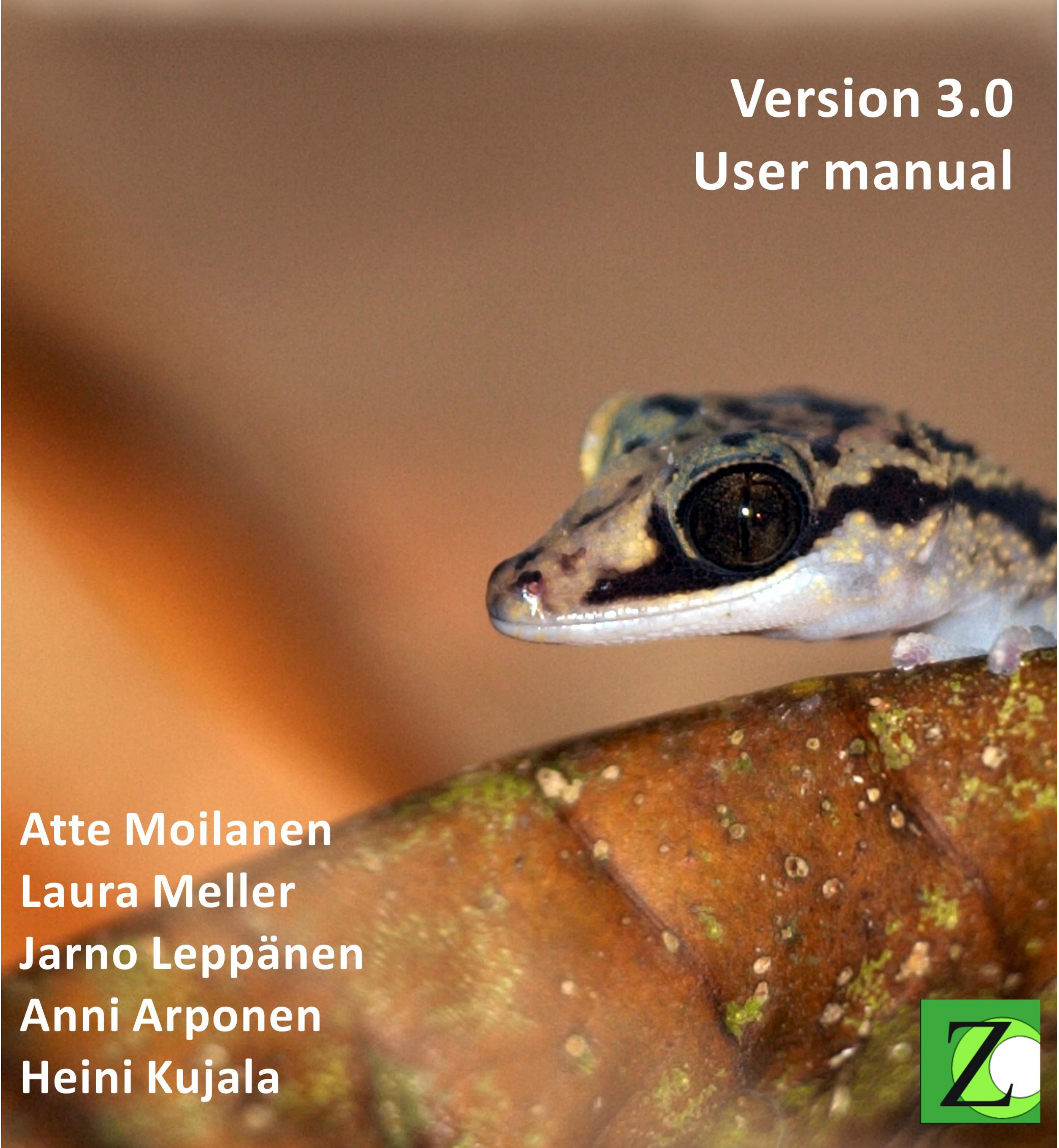
Spatial conservation planning framework and software

ZONATION

Version 3.0
User manual



Atte Moilanen
Laura Meller
Jarno Leppänen
Anni Arponen
Heini Kujala

ZONATION

Conservation planning software

Zonation is a spatial conservation prioritization framework for large-scale conservation planning. It identifies areas, or landscapes, important for retaining high habitat quality and connectivity for multiple biodiversity features (eg. species), providing a quantitative method for enhancing species' long term persistence.

Essentially, this software is a decision support tool for all non-commercial parties working on conservation issues.

As Zonation operates on large grids, it provides a direct link between GIS, statistical distribution modeling and spatial conservation prioritization.

The Zonation framework is presently under constant development and the next version of the software can be expected not too far in the future. Thus, keep an eye on our web site:
www.helsinki.fi/bioscience/ConsPlan

Zonation - User manual

© 2004-2011 Atte Moilanen

All rights reserved. USE THIS SOFTWARE AT YOUR OWN RISK. THE AUTHORS WILL NOT BE LIABLE FOR ANY DIRECT OR INDIRECT DAMAGE OR LOSS CAUSED BY THE USE OR THE INABILITY TO USE THIS SOFTWARE.

CONDITIONS OF USE: Zonation v. 3.0 is freely usable for non-commercial uses. For any other kinds of uses, please contact the author for permission.

Do not use this software if you disagree with the disclaimer or conditions of use. Even though the Zonation software has been done with the best of intentions, it is quite beyond one small research group to ensure its correct operation under all operating systems and environments. Anticipating all potential combinations of erroneous input has not been possible. Therefore, use the software with care and make an effort to understand how the inputs connect to outputs.

Publisher

Biodiversity Conservation
Informatics Group
Department of Biosciences
University of Helsinki, Finland

Credits

Manual

Managing editor
Atte Moilanen

Technical editors
Laura Meller
Anni Arponen
Heini Kujala
Aija Kukkala

Technical content
Atte Moilanen
Jarno Leppänen

Cover Design
Laura Meller
Atte Moilanen

Cover Photo
Evgeniy Meyke

Zonation v.3.0 software

Numerical core
Atte Moilanen
Jarno Leppänen

New GUI
Jarno Leppänen
Atte Moilanen

Thank you

John Leathwick, Alison Cameron, Carlos Carroll, Aldina Franco, Ascelin Gordon, Grzegorz Mikusinski, Bronwyn Rayfield, Chris Thomas, Jim Thomson, and Brendan Wintle, for collaborations and comments on various versions of the software.

Special thanks to Brendan Wintle who generously provided sample data files from the Hunter Valley region to be included with the tutorial, and to Evgeniy Meyke who has kindly provided beautiful background photographs for the Zonation documentation.

Thanks to all who have collaborated in the development of Zonation methods and applications!

Thank you to our supporters:

The Zonation project has been 2004-2011 supported by the Academy of Finland project 1206883 to A.M.; the Ministry of Environment; the Finnish Forest and Park Service; the EU 7th framework project SCALES; EU/ERC project GEDA; and two Academy of Finland Centre of Excellence grants (2000-2005 and 2006-2011) to the Finnish Centre of Excellence in Metapopulation Biology directed by Academy professor Ilkka Hanski.

Helsinki, Nov 22, 2011, Prof. Atte Moilanen
Dept. Biosciences
P.O. Box 65, FI-00014 University of Helsinki
Finland

Table of Contents

Part I Introduction

Aim & purpose 10

The Zonation framework in a nutshell 11

Zonation compared to other reserve selection approaches 12

- Zonation 12
- Integer programming 13
- Stochastic global search 13

A typical Zonation work flow 15

Software installation and quick start 17

New features 19

Part II Methods & algorithms

References 22

The Zonation meta-algorithm 25

Aggregating conservation value: the cell removal rule 27

- Basic core-area Zonation 27
- Additive benefit function 29
- Target-based planning 31
- Generalized benefit function 32
- General differences between cell removal rules 34
- Random removal 37

Inducing reserve network aggregation 38

- Boundary Length Penalty (BLP) 39
- Distribution smoothing 40
- Boundary Quality Penalty (BQP) 41
- Directed connectivity (NQP) 44
- Matrix connectivity 47
- Edge adjustment in connectivity 47

Uncertainty analysis 49

- Uncertainty in species distributions, distribution discounting 50
- Positive uncertainty in species distributions - opportunity analysis 54
- Uncertainty in the effects of landscape fragmentation 54

Ecological interactions 57

Replacement cost analysis 61

Community level analysis 64

Alternative land uses - multi-criterion analysis 67

Landscape condition and retention 69
Landscape dynamics 72
Administrative units 73
Assumptions & limitations 76

Part III ZIG - The Zonation software

Important general information about file operation 79
Running Zonation 80
  Command line 80
  Loading previously calculated Zonation solutions 85
Input files & settings 86
  Introduction 86
  Compulsory files 86
  Biodiversity feature map files 87
  Biodiversity feature list file 90
  Run settings file 95
  Optional files 104
  SSI list and coordinates 104
  Boundary quality penalty definitions file 106
  Directed connectivity layer 108
  Connectivity and community similarity matrices 110
  Connectivity edge effect fix file 111
  Cost layer 112
  Distributional uncertainty map layer 112
  Interactions definition file 115
  Removal mask layer 116
  Analysis area mask 117
  Planning unit layer 118
  Groups file 119
  Alternative land uses layer 121
  Condition layer 121
  Retention layer 122
  Administrative units analysis files 123
  Automated post-processing file 125
Standard Zonation output 129
  Automated file output 129
  Optional output file formats 132
  Output files from optional analyses 133
Post-processing analyses & options 135
  Automated post-processing 135
  Landscape identification 135
  Statistics for management landscapes 136
  Solution comparison 138
  Solution cross-comparison using solution loading 139
What Zonation does NOT do directly 140

Data limitations & system requirements 140
Troubleshooting 143

Part IV Zonation Graphical User Interface

Main menu 148
Project management 149
Visual output 156
Map View 156
Output View 161
Plot View 162

Part V Zonation analysis setups for common planning needs

Basic analysis components 166
Simple Zonation and species weighting 166
BLP Zonation 171
Distribution smoothing + Zonation 172
BQP Zonation 175
Directed freshwater connectivity 178
Matrix connectivity 182
Cost efficiency analysis 184
Analysis with uncertain inputs 186
Ecological interactions 191
Setups and interpretations for basic planning problems 194
Selecting conservation areas 194
Identifying least valuable areas for conservation 196
Expanding conservation areas 198
Evaluating existing/proposed conservation areas 202
Targeting of incentive funding 205
Offsetting and targeting of compensation measures 207
Full analysis setups 209
Community level analysis 209
Combined community and species level analysis 213
Single-species prioritization 215
Balancing alternative land uses considering multiple costs 218
Balancing representation and retention 221
Habitat restoration and dynamic landscapes 225
Setups for climate change 228
Administrative units analysis 232
Setup combinations 236

Part VI Tutorial & Examples

Exercise 1 242
Exercise 2 245

Exercise 3 248
Exercise 4 250
Exercise 5 253
Exercise 6 255
Exercise 7 257
Exercise 8 260
Exercise 9 262
Exercise 10 264
Exercise 11 267

Index

Part I

1 Introduction

1.1 Aim & purpose

Zonation is a framework for conservation prioritization and large-scale spatial conservation planning. It identifies areas, or landscapes, that are important for retaining habitat quality and connectivity simultaneously for multiple species (or any other other biodiversity features), thus providing a quantitative method for enhancing persistence of biodiversity in the long term. Zonation can do traditional reserve selection or site selection as well, but this is only a subset of analyses allowed by conservation prioritization. Typical analyses allowed by Zonation include (i) identification of near-optimal connected reserve networks, (ii) expansion of existing reserve networks, (iii) evaluation of existing or protected reserve networks, (iv) identification of ecologically low-value areas for economic use and (v) prioritization which can be used for many purposes including targeting of incentive funds.

Zonation produces a hierarchical prioritization of the landscape based on the conservation value of sites (cells), accounting for complementarity. The algorithm proceeds by removing least valuable cells from the landscape while minimizing marginal loss of conservation value, accounting for connectivity needs and priorities given for biodiversity features (species, land cover types etc.). As a result, a nested sequence of highly connected landscape structures is obtained with the core areas of species distributions remaining latest and previously-removed areas showing as buffer zones. In this way, landscapes can be zoned according to their potential for conservation, and differing degrees of protection, maintenance or restoration effort can be applied to different zones. The purpose is not necessarily to produce a detailed conservation plan for a large region, but to identify priority areas of the landscape that could be subjected for more detailed analysis and planning that accounts for other land-use pressures than nature conservation.

The Zonation software has been geared towards using large grids as input data, facilitating a direct link between GIS, statistical distribution modeling and Zonation. It is particularly simple to input modeled species distributions, community turnover or land cover types into Zonation. The Zonation software can be run with relatively large datasets on an ordinary desktop PC. Zonation v.3.0. is 64bit software, which allows very large analyses on PCs that have large memory capacity.

The Zonation software is intended for the analysis of biological data with the aim of finding out spatially good conservation solutions. The output of Zonation should be seen as an analysis of conservation value which feeds into a broader land use planning framework where political decisions are made. In other words, Zonation does not construct comprehensive land use plans in that sense. However, commercial or recreational value for alternative land uses can be entered as (opportunity) cost information, thus allowing balancing of alternative land uses. Note that Zonation can also be used for identifying the least important parts of the landscape, those in which human activity would cause least harm to biodiversity value.

1.2 The Zonation framework in a nutshell

Aim and purpose

- To provide a tool for large-scale high-resolution spatial conservation planning using primarily GIS grid data

Analyses

- Identification of optimal conservation areas
- Identification of least useful conservation areas
- Replacement cost analysis for current or proposed reserves
- Balancing of alternative land uses (New in 3.0)
- Combined community-level and species-level prioritization (New in 3.0)
- Planning modes (defining alternative perceptions of conservation value)
    - Core-area Zonation
    - Additive benefit function
    - Target-based benefit function
    - Versatile generalized benefit function
    - Random ranking

Data

- Large-scale grids with
    - Presence/absence -data
    - Probabilities of occurrence
    - Abundance/density -data
    - Uncertainty of occurrence data
    - Cost and mask layers
    - Landscape condition (New in 3.0)
    - Retention of biodiversity (New in 3.0)
- Point observation data
- Planning unit layers
- Administrative units (New in 3.0)
- Community similarity or connectivity matrices (New in 3.0)

Features

- Species priorities via weighting
- Methods for dealing with connectivity needs of species
    - Distribution Smoothing (species-specific)
    - Boundary Quality Penalty (species-specific)
    - Boundary Length Penalty
    - Directed Freshwater Connectivity (species-specific)
    - Interaction connectivity (between two distributions of features)
    - Matrix connectivity (between multiple features; New in 3.0)
    - Edge effect correction (New in 3.0)
- Uncertainty analysis aiming at robust conservation decisions
- Clearly defined trade-offs between species
- Prioritization over multiple administrative regions (New in 3.0)
- Automated post-processing analyses (New in 3.0)
- Utilizing increased memory capacity for versatile analyses (New in 3.0)

1.3 Zonation compared to other reserve selection approaches

In this section we comment on the differences between Zonation and other commonly used approaches to reserve selection. This comparison is not meant to be exhaustive nor completely referenced, but rather, to give an indication of the most fundamental differences (that we believe to exist) between these methods. This section will naturally become outdated when new features are developed into Zonation and other conservation planning software packages.

1.3.1 Zonation

Input data. Zonation is targeted for use with large grid-based data sets. This implies that species distributions, used within Zonation, might be produced using some predictive statistical technique using environmental layers as predictors for species presence/abundance. Data sets in the order of millions of grid cells can be analyzed. Observational data can, of course, be used as well.

Output. Instead of outputting the optimal set of sites for achieving targets, Zonation outputs (1) the hierarchy of cell removal throughout the landscape and (2) feature-specific representation loss curves. This kind of output has multiple advantages.

(i) The result for a range of targets is immediately obvious,
(ii) there is an indication of the importance of all cells, both inside and outside any given fraction,
(iii) the curves show how well (relatively) individual species do at any given fraction of the landscape, and
(iv) the curves indicate the relative value of the solution as well as the stability of the solution.
(v) the zonation output lends itself to easy visualization

We will elaborate the item (iv). If species representation levels are declining rapidly at the chosen landscape fraction, it means that the solution is not stable with respect to uncertainty in input data, and that smallish changes in the selected fraction and / or spatial pattern might have large consequences for the species. If the species performances are stable at the chosen fraction, then small changes in the fraction/spatial pattern are unlikely to have any significant effect on the solution quality.

Additionally, the core-area Zonation method has a specific feature in that it emphasizes best areas for all species, instead of treating low-to-medium quality locations as additive.

Optimality. The optimality characteristics of Zonation have not been conclusively examined, but this is our present evaluation of this issue.

(i) Zonation using additive benefit functions or the targeting benefit function (above the target) is very close to globally optimal. This is because with these cell removal rules the optimization problem is convex and can thus be solved using a gradient-like iterative heuristic (van Teeffelen and Moilanen 2008). Also, with the additive cell removal rules, the degree of sub-optimality goes down when the landscape size (number of cells) increases. Thus optimality is not a problem with the additive cell removal rules. An exception is the use of the Boundary Quality Penalty BQP, which renders the problem non-convex (especially if some species benefit from fragmentation), and the degree of sub-optimality of the solution is unknown.

(ii) The core-area Zonation. This method has so far only been defined algorithmically, not in an objective function form (the CAZ cell removal rule specifies a difference equation for conservation value but not the objective directly), and the degree of sub-optimality of results is unknown. Then again, no other implementation of this method is available.

Limitations. Zonation is primarily intended for binary select or not, protect or not, restore or not -type problems. It is not meant for the direct (near-optimal) targeting of multiple alternative conservation actions, like for example MARXAN with Zones is.

1.3.2 Integer programming

Input data. Can accept arbitrary sites as well as grid cells. According to an older review, Williams et al. (2004), the data size limits of integer programming (IP) were at that time around 10000 landscape elements, which as a grid is only 100x100 elements. While larger data sets can nowadays be processed using IP, Zonation v.3.0. can go up to ~50 million element grids, which is likely to be too much for IP. In addition, some Zonation problem variants are nonlinear (BQP; NQP), and such analyses cannot be approached using IP at all.

Output. Globally optimal set of sites achieving targets. No prioritization through the landscape, no performance curves.

Optimality. Guaranteed globally optimal solution to a simplified problem. The value of the global optimality of results is compromised by the requirement that both the objective function and constraints need to be linear (or that they can be linearized). In a sense you have the optimal solution to the wrong (simplified) problem. Not applicable, at least not easily, to species-specific connectivity calculations on large landscapes (see also Moilanen 2008).

Literature

Williams, J., C.S. ReVelle, and Levin, S.A. 2004. Using mathematical optimization models to design nature reserves. *Frontiers in Ecology and the Environment* 2: 98–105.

Moilanen, A. 2008. Two paths to a suboptimal solution - once more about optimality in reserve selection. *Biological Conservation* 141: 1919-1923.

1.3.3 Stochastic global search

Stochastic global search includes techniques such as simulated annealing (SA; as in MARXAN) and genetic algorithms (Gas).

Input data. In principle can be run on extremely large problems with few constraints on the complexity of the problem. SA can handle larger problems than a GA, because of the memory requirement for storing the GA population.

Output. A solution to the problem, typically of unknown quality. In some cases it may be possible to devise an analytical method that provides bounds on solution quality (as in Moilanen 2005), which then changes the method from a heuristic to an approximation. (Heuristic = method for which the quality of results is unknown; approximation = method for which the maximum degree of sub-optimality of the results has been quantified in a non-trivial manner.)

Optimality. Degree of sub-optimality will be highly dependent on (1) the size of the data, (2) the complexity of the problem, like does it have nonlinear connectivity components in it, and (3) the details of the implementation of the optimization algorithm. SA and GA are by no means standard algorithms (except for the high-level meta-algorithm). They can be varied in endless ways, in particular, in terms of how they generate the new solutions to evaluate. If the search starts far from the good regions of the search space, it actually is not guaranteed that the good regions are found at all. Good convergence with large problems is absolutely not guaranteed. Multiple runs from different starting points are required to test for indications of convergence – and if multiple runs reliably converge to a very similar result, then this indeed is an indication that the solution probably is quite acceptable in terms of optimality. Probably the situation is rather good with smallish data sets with thousands or tens of thousands of sites, but at the million-element scale the performance of these methods is poorly known. Relative performance probably degrades when problem size increases, which is opposite from what is actually expected for Zonation, at least with the additive cell removal rules.

There are piles of literature on optimization, which is an enormous field of science in itself. See the references below for examples of the use of stochastic optimization on nonlinear reserve selection problems. Also check MARXAN reserve selection software user manual and references therein.

Literature

Moilanen, A. 2005. Reserve selection using nonlinear species distribution models. *American Naturalist* 165: 695-706. AND in particular its electronic appendixes A-C.

Moilanen, A. and M. Cabeza. 2002. Single-species dynamic site selection. *Ecological Applications* 12: 913-926.

Moilanen, A. and I. Ball. 2009. Heuristic and approximate optimization methods for spatial conservation prioritization. Pp. 58-69 in Moilanen, Wilson and Possingham (Eds.), *Spatial Conservation Prioritization*, Oxford University Press.

1.4 A typical Zonation work flow

This section outlines a typical sequence of steps that would be done for the Zonation analysis of one data set.

(i) Get the basic analyses running

i.1. Install Zonation, and make sure you get the program running with the example files provided.

i.2. Decide your cell removal rule.

i.3. Produce a new settings file, species list file etc. for your own data and check that you are able to run the basic analysis (without aggregation methods or uncertainty analysis).

i.4. Try variants of the basic analysis by adding unequal species weights, aggregation methods, uncertainty analysis and some other potentially relevant analysis features. You can use solution comparison to check how big a difference does the addition of one complication cause into the solution. These preliminary analyses can be run using high warp factors (100-1000) to reduce runtimes.

(ii) Identify your base-analysis. Zonation v.3.0. provides a multitude of options about which considerations to include in your analysis. All of them have some specific data requirements, and applying them all is hardly a meaningful thing to do. Your choices of analysis features would depend on your specific planning context, objectives and data availability. Further, there are endless options of how to weight species or other biodiversity features, which aggregation method to use with what exact parameter values, and defining parameter values for other analysis features. You cannot run all combinations of everything and indeed it is not useful to do so. Therefore, after getting the basic Zonation running, you need to decide the most reasonable options for your analysis. Things that need to be decided include

ii.1. Decide the most appropriate analysis setup to match your needs. Analysis setups for different planning problems are outlined in section 5. Decide what would be the best combination of analysis components: should you consider cost, uncertainty, interactions, landscape condition and retention, alternative land uses, landscape dynamics etc.

ii.2. Decide about how to induce aggregation into the final solution. Options include distribution smoothing, boundary quality penalty, directed freshwater connectivity, boundary length penalty and matrix connectivity. In general, you want aggregation at least if your planning units are small, around hectares or so, because with small selection units population dynamics of nearby cells are strongly linked. If planning units are very large, e.g. 10x10km cells, then aggregation could plausibly be omitted.

ii.3. Decide weights for species or other biodiversity features. Equal weights is the default option but there may well be good reason to favor particular features by giving them more weight.

(iii) Base-analysis and sensitivity analysis. At this point you have identified the analysis options which you believe to be most appropriate. Next iii.1 Run your base-analysis, preferably using a relatively low warp factor.

iii.2 Run variants around your base-analysis varying a single analysis feature at a time (you probably cannot run all combinations of everything). This is essentially a sensitivity analysis, which is done by varying weights, aggregation and uncertainty analysis settings within reasonable bounds. An advisable strategy for complex analysis setups is to start from a simple setup and adding one new component at a time, always checking the impact of that component. Use solution comparison to see how big a difference various options make.

iii.3 An analysis of selection frequency may provide useful summary information over analyses. This option will be added to the upcoming version of Zonation v.3.0 GUI.

At this point you have a good idea of how different planning options influence your analysis and solutions.

(iv) Interpretation and post-processing of output. Identify priority areas for conservation (or the least important areas) from the solution. Identify management landscapes and check their statistics to find out why different areas are important – what are the biodiversity features that occur there?

(v) Evaluation of proposed reserve areas using replacement cost analysis. If you need to evaluate proposed or existing reserve areas, you can do that using mask files and replacement cost analysis. This involves repeating your base analysis both with and without existing/proposed areas included/excluded.

1.5 Software installation and quick start

Installation

The installation package includes the Zonation program (**zig3.exe**), the user interface (**zig3gui.exe**), a user manual (**.pdf**) and tutorial files. In addition, it contains the following DLL libraries which need to be placed in the same directory as zig3.exe for it to work properly: gdal18.dll, libfftw3f-3.dll, msvcp90.dll, msvcr90.dll, QtCore4.dll, QtGui4.dll, QtNetwork4.dll, qwt5.dll and zig3lib.dll.

Because with Zonation v.3.0 the DLL libraries need to be located within the same directory as the .exe, it is no longer recommendable to make a new copy of the program to the directory containing data files when starting a new project. Instead, it is more convenient to establish a permanent directory for the executable and the DLL files and to call the program from that directory, e.g. "call C:\Zonation\zig3.exe". For simplicity, the example calls in the manual do not contain a file path.

Quick start

Here are instructions to run the basic Zonation for those who have already familiarized themselves, at least to some extent, with the program. For more detailed instructions and additional analyses please see sections 3.2. Running Zonation, 3.3. Input files & settings, 3.5. Post-processing analyses & options, 4. Zonation Graphical User Interface and 5. Analysis setups for common planning needs.

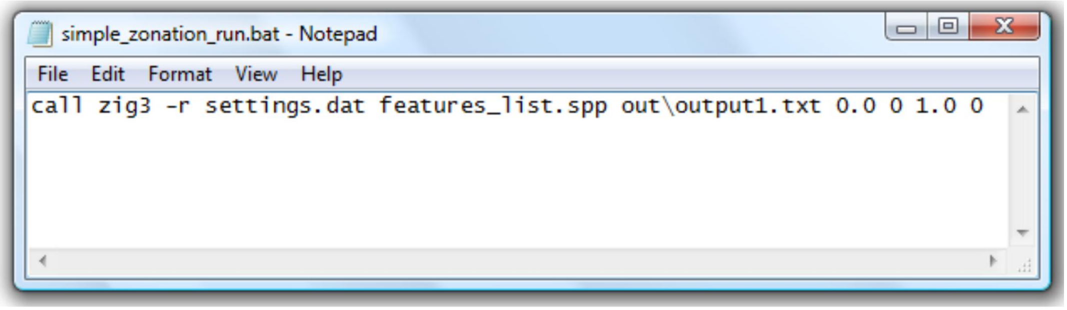
1. To run the program you need at least the following **input files**:
    - Biodiversity feature map files, which are basic raster files (.asc or .img or .tif) exported from GIS programs. These files define species distributions in the landscape. The program can incorporate any kind of species distribution data, such as presence-absence, probabilistic or abundance data, or species-specific population connectivity surfaces etc., as longs as data for all species is in the same format. The data can also be habitat or community types. Zonation can also use point distribution data.
    - The names of all the biodiversity feature files must be listed in a separate biodiversity feature list file (.spp), each file on a separate row with the species-specific parameters before the file name (see section 3.3.2.2. for more detailed descriptions). The species list file tells the program which species distribution files will be used in the analysis.
    - A run settings file to define the settings in your analysis
    - A batch file containing the command line.

Remember to always use decimal points, NOT commas, in all input files!

3. Select the suitable cell removal rule in your run settings file.

4. Type the following command line to notepad:
```
call zig3 -r settingsfile.dat specieslistfile.spp
outputfile.txt 0.0 0 1.0 0
```
and save the file in .bat format into the same directory where you have all the other files.

In this command line, give the names of your settings file and species list file and define a suitable name for your output files. See section 3.2.1. for explanations for the four numbers in the call.



```
call zig3 -r settings.dat features_list.spp out\output1.txt 0.0 0 1.0 0
```

5. Open the Zonation GUI (**zig3gui.exe**). Right-click your mouse on the **Project View** and use the dialogue window for selecting the right batch file. Initialize the Zonation run by right-clicking the .bat file name in the Project View and selecting **Queue**.

6. Monitor your run by double-clicking the run in the GUI **Process View**. Output map describing the cell ranking will be drawn in the **Map View**. Representation curves are drawn in the **Plot View**. A memo of the run is written in the **Output View**. See more detailed description of visual output in section 4.3.

The program automatically produces eight output files with all analysis variants:

- .jpg and .emf maps of the landscape ranking showing the order of cell removal in different colors. See section 4.3.1. for detailed interpretation of the colors.
- A .curves.txt -text file containing a list of species and weights used in the analysis, and columns representing how large proportion of distribution of each species is remaining when landscape is iteratively removed.
- A .rank.asc -raster file representing the order of cell removal (ranking). This file can be used to produce map images in GIS softwares.
- A .prop.asc -raster file representing proportions of species distributions (across species) remaining at the removal of that cell. This file can be used to produce map images in GIS softwares.
- A .wrscr.asc -raster file. This file contains a weighted range size normalized measure of conservation value for each cell, which can be used as a scoring measure of value for cells.
- A .txt file. This file is written during the run. It keeps track of the files, settings and analyses used during the runs and is useful for tracing errors and checking that everything happened as supposed to. Note that some error messages or warnings may appear here.
- A .run_info.txt -text file keeping track of input files, analyses and settings used. This file will be created after you have closed the program.

Depending on analysis variant, there may be additional output files.

1.6 New features

This section shortly lists the new features and small additions that are included in Zonation v. 3.0 in comparison to earlier versions. It also presents some useful tricks you can use with Zonation.

Added for v. 3.0

- **Graphical User Interface**. Zonation v.3.0 has a completely new GUI, mostly designed and implemented by Jarno Leppänen. The new GUI includes a facility for administrating Zonation projects, which may consist of multiple runs. The GUI also includes facilities for examining and editing input files, visualizing and editing output maps, and for checking of the consistency of input file sets (section 4).
- **Matrix connectivity** to allow accounting for connectivity to multiple features simultaneously. Decreased matrix connectivity of cells on the edges of the planning region is corrected with the edge correction operation (sections 2.3.5. and 5.1.5.).
- **Removal mask layer** has a new format and improved function. Multiple levels of prioritization hierarchy are now possible (section 3.3.3.9.).
- **Analysis area mask** with which you can select a subset of cells within your planning region to be included in the analysis (section 3.3.3.10.).
- **Alternative land-uses** to allow consideration of multiple land uses to alleviate conflicting interests – in practice, the value grids are input as biodiversity feature files, and assigned negative weights (sections 2.8. and 5.3.4.).
- **Landscape condition and retention** (sections 2.9. and 5.3.5.)
- **Community level analysis** with the focus on community composition. Overlap in species composition is accounted for by similarity expansion (sections 2.7. and 5.3.1.).
- **Administrative units** to allow planning over multiple administrative regions (sections 2.11. and 5.3.9.).
- **Automated post-processing analyses** for identifying management landscapes and comparing two solutions in terms of overlap and cell removal order (section 3.5.1.).
- **Utilizing the increased memory capacity for versatile analyses**
- **New output** Zonation v. 3.0 has new outputs: a .log file for tracking error messages and warnings (see section 3.4.2), an .emf map instead of .bmp, and a few new output files specific to some new optional analyses (see section 3.5.3.). In addition, output maps can be produced in several raster and image formats (see section 3.4.2.)
- **Additional info in the Memo -window** Because of all the cool features listed above, there is a lot of new information printed into the Memo/run info.

Outlook for the next version

Resources for the development of Zonation have been secured for at least the next couple of years. While specific details of developments are unknown at the present time, a major software update can be expected in 2-3 years.

Part II

2 Methods & algorithms

2.1 References

The basic Zonation analysis and distribution smoothing

Moilanen, A., Franco, A. M. A., Early, R., Fox, R., Wintle, B., and C. D. Thomas. 2005. Prioritising multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society of London, Series B, Biological Sciences* 272: 1885-1891.

Moilanen, A. 2007. Landscape zonation, benefit functions and target-based planning: Unifying reserve selection strategies. *Biological Conservation*, 134: 571-579.

Distribution smoothing and analysis of surrogacy power
Franco, A.M.A., Anderson, B., Roy, D.B., Gillings, S., Fox,R., Moilanen, A., and C.D. Thomas. 2009. Surrogacy and persistence in reserve selection: landscape prioritisation for multiple taxa in Britain. *Journal of Applied Ecology*, 46: 82-91.

Distribution smoothing, info-gap uncertainty analysis
Moilanen, A. and B. A. Wintle. 2006. Uncertainty analysis favours selection of spatially aggregated reserve structures. *Biological Conservation,* 129: 427-434.

Basics of the information-gap decision theory for reserve selection uncertainty analysis
Moilanen, A., Runge, M. C., Elith, J., Tyre, A., Carmel, Y., Fegraus, E., Wintle, B., Burgman, M. and Y. Ben-Haim. 2006a. Planning for robust reserve networks using uncertainty analysis. *Ecological Modelling,* 199 (1): 115-124.

Accounting for distributional uncertainty in Zonation, the distribution discounting method
Moilanen, A., B. A. Wintle, J. Elith and M. Burgman. 2006b. Uncertainty analysis for regional-scale reserve selection. *Conservation Biology,* 20: 1688-1697.

A quantitative method for generating reserve network aggregation
Moilanen, A., and B. A. Wintle. 2007. The boundary-quality penalty: a quantitative method for approximating species responses to fragmentation in reserve selection. *Conservation Biology*, 21: 355-364.

Connectivity

An extension of the BQP method to freshwater systems with different connectivity requirements upstream and downstream
Moilanen, A., Leathwick, J. and J. Elith. 2008. A method for freshwater conservation prioritization. *Freshwater Biology*, 53: 577-592.

Matrix connectivity in expanding current protected area network
Lehtomäki, J., Tomppo, E., Kuokkanen, P. Hanski, I., and A. Moilanen. 2009. Planning of forest conservation areas using high-resolution GIS data and software for spatial

conservation prioritization. *Forest Ecology and Management*, 258: 2439-2449.

Single-species conservation prioritization with multiple connectivity considerations

Rayfield, B., Moilanen, A. and M.-J. Fortin. 2009. Incorporating consumer-resource spatial interactions in reserve design. *Ecological Modelling*, 220: 725-733.

Edge adjustment for matrix connectivity

Arponen, A., Lehtomäki, J., Leppänen, J., Tomppo, E. and Moilanen, A. Submitted manuscript. Analysis resolution and connectivity in large-scale spatial conservation prioritization.

Benefit function approach to reserve selection

Arponen, A., Heikkinen, R., Thomas, C.D. and A. Moilanen. 2005. The value of biodiversity in reserve selection: representation, species weighting and benefit functions. *Conservation Biology*, 19: 2009-2014.

Arponen, A., Kondelin, H. and A. Moilanen. 2007. Area-Based Refinement for Selection of Reserve Sites with the Benefit-Function Approach. *Conservation Biology*, 21 (2): 527–533.

Moilanen, A. and M. Cabeza. 2007. Accounting for habitat loss rates in sequential reserve selection: simple methods for large problems. *Biological Conservation*, 136: 470-482.

van Teeffelen, A., and A. Moilanen. 2008. Where and how to manage: Optimal allocation of alternative conservation management actions. *Biodiversity Informatics*, 5: 1-13.

Replacement cost analysis

Cabeza, M. and A. Moilanen. 2006. Replacement cost: a useful measure of site value for conservation planning. *Biological Conservation,* 132: 336-342.

Moilanen, A., Arponen, A., Stockland, J.N. & M. Cabeza. 2009. Assessing replacement cost of conservation areas: How does habitat loss influence priorities? *Biological Conservation,* 142: 575-585.

A replacement cost analysis and consideration of cross-taxon surrogacy in conservation planning

Kremen, C., A. Cameron, A. Moilanen, S. Phillips, C. D. Thomas, H. Beentje, J. Dransfeld, B. L. Fisher, F. Glaw, T. Good, G. Harper, R.J. Hijmans, D. C. Lees, E. Louis Jr., R. A. Nussbaum, A. Razafimpahanana, C. Raxworthy, G. Schatz, M. Vences, D. R.Vieites, P. C. Wright and M. L. Zjhra. 2008. Aligning conservation priorities across taxa in Madagascar, a biodiversity hotspot, with high-resolution planning tools. *Science*, 320: 222-226.

A replacement cost analysis with several levels of financial cost constraints applied to marine protected areas

Leathwick, J. R., Moilanen, A., Francis, M., Elith, J., Taylor, P. Julian, K. and T. Hastie. 2008. Novel methods for the design and evaluation of marine protected areas in offshore waters. *Conservation Letters*, 1: 91-102.

Other relevant references

Community-level conservation prioritization
Arponen, A., Moilanen, A., and S. Ferrier. 2008. A successful community-level strategy for conservation prioritization. *Journal of Applied Ecology*, 45: 1436-1445.

Community-level conservation prioritization and directed connectivity
Leathwick, J.R., A. Moilanen, S. Ferrier and K. Julian. 2010. Complementarity-based conservation prioritization using a community classification, and its application to riverine ecosystems. *Biological Conservation*, 143: 984-991.

Moilanen, A., Anderson, B.J., Eigenbrod, F., Heinemeyer, A., Roy, D.B., Gillings, S., Armsworth, P.R., Gaston, K.J. and Thomas, C.D. 2011. Balancing alternative land uses in conservation prioritization. *Ecological Applications*, 21: 1419-1426.

Balancing representation and landscape retention in community-level conservation prioritization
Moilanen, A., Leathwick, J.R. and J.M. Quinn. 2011. Spatial prioritization of conservation management. *Conservation Letters*, 4: 383-393.

Habitat restoration and landscape dynamics
Thomson, J.R., Moilanen, A., McNally, R., and P. Vesk. 2009. Where and when to revegetate: A quantitative method for scheduling landscape reconstruction. *Ecological Applications*, 19: 817-828.

Utilizing ecological interactions in Zonation to account for climate change in conservation planning
Carroll, C., Moilanen, A., and J. Dunk. 2010. Designing multi-species reserve networks for resilience to climate change: priority areas for spotted owl and localized endemics in the pacific North-West USA. *Global Change Biology*, 16: 891-904.

Balancing local and global representation over multiple administrative regions
Moilanen, A. and A. Arponen. 2011. Administrative regions in conservation: Balancing local priorities with regional
to global preferences in spatial planning. *Biological Conservation*, 144: 1719-1725.

Identifying least suitable areas for conservation to expand urban areas
Gordon, A., Simondson, D., White, M., Moilanen, A., and Bekessy, S.A. 2009. Integrating conservation planning and land-use planning in urban landscapes. *Landscape and Urban Planning*, 91: 183-194.

Z-relevant Book Chapters

Moilanen, A. and I. Ball. 2009. Heuristic and approximate optimization methods for spatial conservation prioritization. Pp. 58-69 in Moilanen, Wilson and Possingham (Eds.), Spatial Conservation Prioritization, Oxford University Press.

Moilanen, A., Kujala, H. and J. Leathwick. 2009. The Zonation framework and software for conservation prioritization. Pp. 196-210 in Moilanen, Wilson and Possingham (Eds.), Spatial Conservation Prioritization, Oxford University Press.

The manuals

This version: www.helsinki.fi/bioscience/consplan

Moilanen, A. and H. Kujala. 2008. Zonation spatial conservation planning framework and software v. 2.0, User manual, 136 pp.

Moilanen, A. and H. Kujala. 2006. Zonation spatial conservation planning framework and software v. 1.0, User manual, 126 pp. Edita, Helsinki, Finland.

For those who would wish to familiarize themselves more broadly with recent literature concerning spatial conservation planning, we recommend using Web of Science (or a similar search facility) with key words such as, **reserve selection**, **reserve network design**, **site selection algorithm**, **area prioritization**, **spatial conservation planning** and **spatial optimization**. Journals such as *Biological Conservation*, *Conservation Biology* , *Ecological Applications*, *Journal of Applied Ecology* and *Environmental Modeling and Assessment*, among others, include many studies concerning quantitative conservation prioritization methods.

2.2 The Zonation meta-algorithm

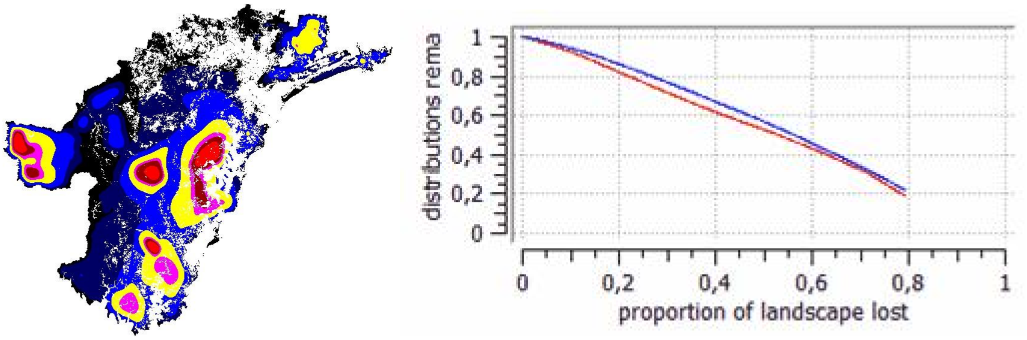
The Zonation algorithm (Moilanen et al. 2005) produces a hierarchical prioritization of the conservation value of a landscape, hierarchical meaning that the most valuable 5% is within the most valuable 10%, the top 2% is in the top 5% and so on. Generally speaking, Zonation simply iteratively removes cells one by one from the landscape, using minimization of marginal loss as the criterion to decide which cell is removed next. The order of cell removal is recorded and it can later be used to select any given top fraction, for example best 10%, of the landscape. Simultaneously, information is collected about the decline of representation levels of species during cell removal.

Essentially, the algorithm applied by Zonation is a reverse, accelerated, iterative heuristic. Reverse comes from starting from the full landscape and removing cells (this is important for the treatment of connectivity). Accelerated comes from the option of removing more than one cell at a time, via the warp factor (see section 3.3.2.3.).

The Zonation meta-algorithm

1. Start from the full landscape. Set rank $r = 1$.
2. Calculate marginal loss following from the removal of each remaining site i, δ_i. Complementarity is accounted for in this step.
3. Remove the cell with smallest δ_i, set removal rank of i to be r, set $r=r+1$, and return to 2 if there are any cells remaining in the landscape.

Thus, sites are ranked based on biological value and the least valuable cells are removed one (or more) at a time, producing a sequence of landscape structures with increasingly important biodiversity features remaining. We want to emphasize that the result of a Zonation analysis is not a single set of areas. Rather, it is (i) the nested ranking of cells and (ii) a set of performance curves describing the performance of the solution at the given level of cell removal.



Example figures illustrating the ranking and the curves.

The Zonation meta-algorithm can, among other things, answer two questions frequently encountered in conservation biology:

- which parts of the landscape, totalling *x*% of landscape cost or area, have the highest conservation priority (ranking), or,
- which part of the landscape includes at least *y*% of the distribution of each species (proportional coverage selection)?

Whether the Zonation algorithm makes any sense at all depends on the definition of marginal loss (δi), step 2 in the algorithm above. This definition is done by a separate cell removal rule (section 2.3.), which implements our conception of how conservation value is aggregated across the landscape and across species (or whatever features). The Zonation method can thus be divided into two parts, the Zonation meta-algorithm and the cell removal rule (= definition of marginal loss), which should not be confounded. The cell removal rule should be seen as a separate component with several alternatives that have different interpretations. Note that the notion of complementarity is inherent in the way the cell removal rule is defined.

There is one feature, which according to Moilanen et al. (2005) is a part of the Zonation algorithm, but which is more appropriately seen as a relevant detail for which there are alternatives. This is edge removal (see section 3.3.2.3.), a feature that allows cells to be removed only from the edge of the remaining landscape. Edge removal promotes maintenance of structural habitat continuity in the removal process. It also makes the cell removal process much faster with large landscapes, which is the primary reason for using it.

Literature

Moilanen, A., Franco, A. M. A., Early, R., Fox, R., Wintle, B., and Thomas, C.D. 2005. Prioritising multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society of London, Series B, Biological Sciences* 272: 1885-1891.

2.3 Aggregating conservation value: the cell removal rule

This section is mainly based on Moilanen (2007).

The Zonation meta-algorithm is the same for all analyses described in this manual. The actual removal order of cells, then, depends on the cell removal rule. The rule determines which cell leads to smallest marginal loss of biodiversity value. There are three conceptually different cell removal rules:

1. Core-area Zonation (section 2.3.1.)
2. Additive benefit function (section 2.3.2.)
3. Target-based planning (section 2.3.3.)
4. Generalized benefit function (section 2.3.4.), a two-piece power function that can assume very versatile forms, allowing flexibility in the specification of conservation value.

Sometimes it may be necessary to create networks where sites are removed in a random order. In order to, for example, compare different methods and their effectiveness in prioritization, one may want to do randomized prioritizations to get a baseline representation level. This is possible in Zonation with the fifth cell removal rule,

5. Random removal (section 2.3.6.).

Note that core-area Zonation has the property that it can identify important, otherwise species-poor, locations where a single species (or few species) has an important occurrence. The additive benefit function analysis gives more weight to locations with high species richness. Therefore, it may be useful to run both analyses and compare the results. If the top-fractions do not agree, then there are some species-rich areas but also some species-poor areas with occurrences of otherwise rare species. Thus running both core-area Zonation and the additive benefit function analysis may reveal information that is interesting for conservation planning.

Literature

Moilanen, A. 2007. Landscape zonation, benefit functions and target-based planning: Unifying reserve selection strategies. *Biological Conservation*, 134: 571-579.

2.3.1 Basic core-area Zonation

This section is mainly based on Moilanen *et al.* (2005) and Moilanen (2007).

In basic core-area Zonation cell removal is done in a manner that minimizes biological loss by picking cell i that has the smallest value for the most valuable occurrence over all species in the cell. In other words, the cell gets high value if even one species has a relatively important occurrence there. The removal is done by calculating a removal index δ_i (minimum marginal loss of biological value) for each of the cells, where:

$$\delta_i = \max_j \frac{q_{ij} w_j}{c_i}, \qquad (1a)$$

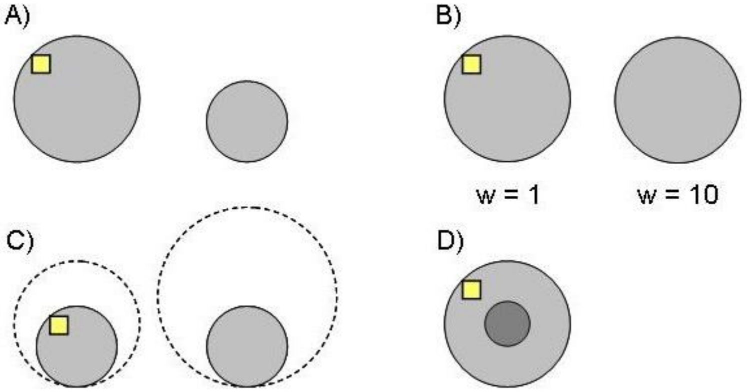
where w_j is the weight (or priority) of species *j* and c_i is the cost of adding cell *i* to the reserve network. When running the analysis, the program goes through all cells and calculates them a value δ_i based on that species that has the highest proportion of distribution remaining in the specific cell (and thus represents the highest biological value to be lost, if the cell is removed). The cell which has the lowest δ_i -value, will be removed.

The critical part of the equation is q_{ij}, the proportion of the remaining distribution of species *j* located in cell *i* for a given set of sites (the set of cells remaining). When a part of the distribution of a species is removed, the proportion located in each remaining cell goes up. This means that Zonation tries to retain core areas of all species until the end of cell removal even if the species is initially widespread and common. Thus, at first only cells with occurrences of common species are removed. Gradually, the initially common species become more rare, and cells with increasingly rare species occurrences start disappearing. The last site to remain in the landscape is the cell with the highest (weighted) richness. This is the site that would be kept last if all else was to be lost. Note that Eq. (1a) can alternatively be expressed as (Moilanen et al. 2005)

$$\delta_i = \frac{1}{c_i} \max_j \frac{q_{ij} w_j}{q_j}, \qquad (1b)$$

where q_{ij} is the fraction of the original full distribution of species *j* residing in cell *i* according to data, and q_j is the fraction of the original distribution of species *j* in the remaining set of cells.

The min-max structure of the equation also indicates a strong preference to retaining the best locations with highest occurrence levels. Thus, the program can spare otherwise species-poor cells, if they have a very high occurrence level for one rare species. It is important to understand that core-area Zonation does not treat probabilities of occurrence as additive; ten locations with p=0.099 is not the same as one location with p=0.99. However, this is strictly true only when analysis is based on biological value only and when a landscape cost layer is not used in the analysis. When cost is used, cell removal is based on local conservation value divided by cell cost (efficiency), and now a high value for a cell can be explained with either (i) a very high occurrence level for some species or (ii) low cost for the cell. Thus, when cost information is used, the interpretation of a core-area becomes vague, and this should be recognized in planning. Therefore it is not recommended to use cost layers when trying to find out biologically most important areas with core-area Zonation.



This figure illustrates principles that core-area Zonation implements in numerical form. Essentially, the question is, if you have two (multiple) species, and you are going to lose a fraction (here one cell, marked as yellow) of one distribution, then where would you prefer to lose the cell from? (A) If you have two otherwise identical species, but one has a larger range remaining, then you prefer to lose from the species that has the larger range. (B) If you have two otherwise equal species, but one has relatively higher weight, then you prefer to lose from the distribution of the species with a lower weight. (C) You have two presently equal species with equally wide distributions. Then you prefer to lose from the species that has had a smaller historical reduction in the range (dashed line). (D) Within the distribution of a species, one prefers to lose from a location with a relatively low occurrence density (light gray).

Literature

Moilanen, A. 2007. Landscape zonation, benefit functions and target-based planning: Unifying reserve selection strategies. *Biological Conservation*, 134: 571-579.

Moilanen, A., Franco, A. M. A., Early, R., Fox, R., Wintle, B., and Thomas, C.D. 2005. Prioritising multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society of London, Series B, Biological Sciences* 272: 1885-1891.

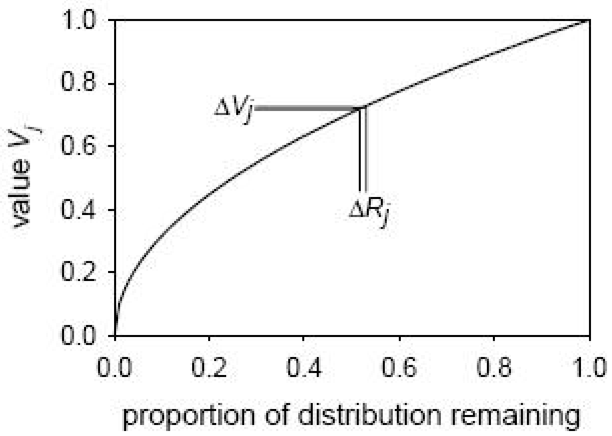
2.3.2 Additive benefit function

Additive benefit function was described by Moilanen (2007).

Compared to core-area Zonation, the additive benefit function takes into account all species proportions in a given cell instead of the one species that has the highest value. The program calculates first the loss of representation for each species as cell *i* is removed, and the δ_i -value of the cell is simply the sum over species-specific declines in value following the loss of cell *i*:

$$\delta_i = \frac{1}{c_i} w_j \sum_j \Delta V_j = \frac{1}{c_i} w_j \sum_j \left[V_j(q_j) - V_j(q_j - i) \right],$$

in which q_j is the representation of species *j* in remaining set of sites, and (q_j-*i*) indicates the set of remaining cells minus cell *i*. Here w_j is the weight of the species *j* and c_i is the cost (or area) of planning unit *i*. Again the cell that has the smallest δ -value, will be removed.



Above is a picture of a benefit function for species *j*. When a grid cell is removed from the landscape, the representation of each species occurring in the removed cell goes down by a small fraction ΔR_j and the respective value for that species goes down by ΔV_j. The total marginal loss in value is simply a sum over species-specific losses. Note that here the species has a standard weight of 1.0, but as with core-area Zonation it is possible to weight species differently when using additive benefit function. The effects of weighting are seen on the scale of the y-axis which will go from 0.0 to species weight w_j instead of going from 0.0 to 1.0.

Because the additive benefit function sums value over all species, the number of species in a cell has a higher significance compared to basic core-area Zonation. For example, using additive benefit function might lead to situation where species-poor cells are removed even though they have a high occurrence level for one or two rare species, because the δ_i -value of these cells is smaller than that for cells that have many common species with high representations. Thus, using the additive benefit function typically results in a reserve network that has a higher performance on average over all species, but which retains a lower minimum proportion of original distributions for the worst-off species compared to core-area Zonation (see figure of the first three removal rules in section 2.3.5 .).

Literature

To find more information about the use of benefit functions, see

Arponen, A., Heikkinen, R., Thomas, C.D. and A. Moilanen. 2005. The value of biodiversity in reserve selection: representation, species weighting and benefit functions. *Conservation Biology* 19: 2009-2014.

Arponen, A., Kondelin, H. and Moilanen, A. 2006. Area-based refinement for selection of reserve sites with the benefit-function approach. *Conservation Biology*, 21: 527-533.

Cabeza, M. and A. Moilanen. 2006. Replacement cost: a useful measure of site value for conservation planning. *Biological Conservation*,132: 336-342.

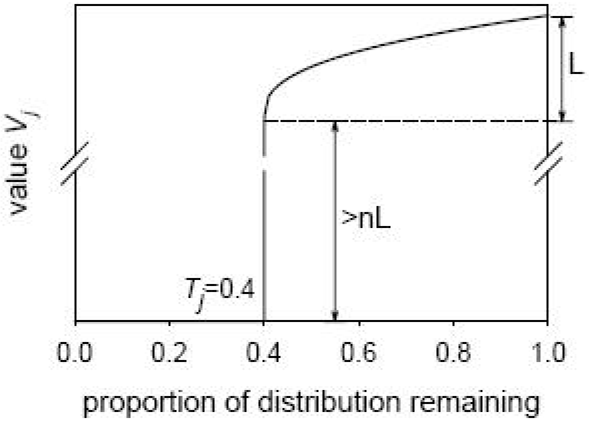
Moilanen, A. 2007. Landscape zonation, benefit functions and target-based planning: Unifying reserve selection strategies. *Biological Conservation*, 134: 571-579.

Moilanen, A. and M. Cabeza. 2007. Accounting for habitat loss rates in sequential reserve selection: simple methods for large problems. *Biological Conservation*, 136: 470-482.

van Teeffelen, A., and A. Moilanen. 2008. Where and how to manage: Optimal allocation of alternative conservation management actions. *Biodiversity Informatics*, 5: 1-13.

2.3.3 Target-based planning

Target-based planning is implemented in Zonation by using a very particular type of a benefit function - the purpose of this special functional form is to enable the Zonation process to converge to a solution that is close to the proportional coverage minimum set solution for the data. In this function value V_j is zero until representation R_j reaches the target T_j. Then there is a step with the height of (n+1), where n is the number of species. When R_j increases above T_j and approaches 1, there is a convex increase in value, with a difference in value $[V_j(1)-V_j(T_j)]=1$. This means that the loss in value from dropping any one species below the target is higher than any summed loss over multiple species that stay above the target.



The idea is that, as cells are iteratively removed, species representations will approach the species-specific targets from above, and that the convex formulation with increasing marginal losses will force species to approach targets in synchrony in terms of lost value. Thus, as one of the species approaches the target level, the program starts to avoid removing cells that contain that particular species (at the expense of other species) in order to retain the target. At some point it will not be possible to remove any more cells without violating the target for at least one species. After one of the species has declined below target, the remaining distribution of that species has no value for the reserve network. Thus removing cells where only this species occurs does not increase the loss of biological value from network anymore.

Note however, that the definition of how marginal value is calculated does not change from that of additive benefit function (section 2.3.2.). Also with this cell removal rule species occurrences are considered as additive and the cell that has the lowest marginal value summed across all species will be removed next.

Also, when using target-based planning the species-specific weights have no function as the goal is to retain a given proportion of distributions for all of the species. However, it is possible to set different targets to different species. It is also recommended to avoid using very high warp factors to allow the program to find the most optimal solution near the targets.

A solution computed with the target-based planning cell removal rule needs to be interpreted with special care. For the other cell removal rules, conservation value from the perspective of one species increases as the number of cells in the landscape decreases. With target-based planning, the conservation value for one species in all cells goes down to zero immediately after the target has been violated. Using targets often leads to a non-optimal solution, especially at resource levels smaller than what allows achieving all targets. If you are using the target-based planning removal rule, we recommend running the same analysis also with another cell removal rule so that you can evaluate the sensitivity of the solution to using targets.

Literature

Moilanen, A. 2007. Landscape zonation, benefit functions and target-based planning: Unifying reserve selection strategies. *Biological Conservation*, 134: 571-579.

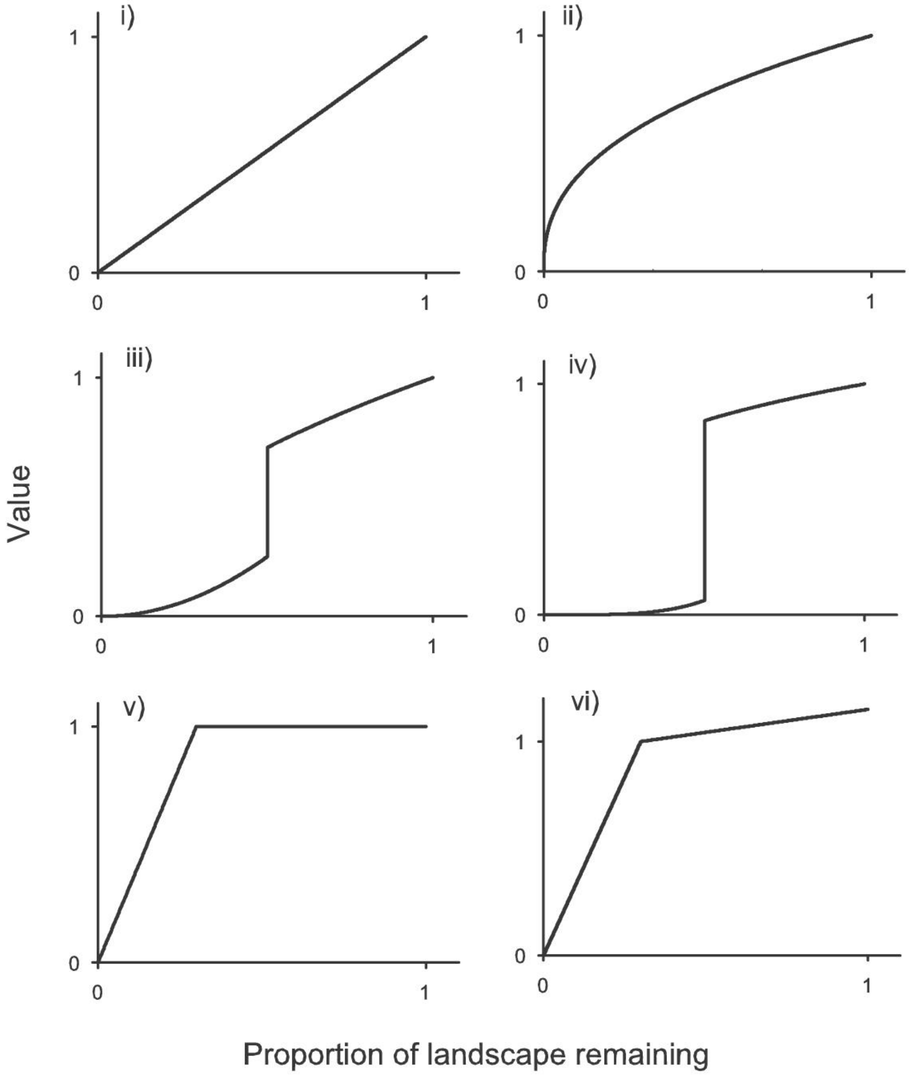
2.3.4 Generalized benefit function

Cell removal rule number four is a generalized benefit function form that allows very flexible function shapes. The function is defined in two pieces, each a power function.

$$V_j(R_j)) = \left\{ \begin{array}{ll} w_1 \left(\frac{R_j}{T_j} \right)^x & \text{if } R_j \le T_j \\ w_1 + w_2 \left(\frac{R_j - T_j}{1 - T_j} \right)^y & \text{if } T_j < R_j < 1 \end{array} \right\}$$

In this equation R_j is the fractional representation level of the species, the fraction of the original distribution remaining. T_j is a nominal target level for the species, at $R_j = T_j$, the representation of the species gets value w_1. x is the parameter of the first part of the power function. When $T_j < R_j < 1$ the function continues as another power function with parameter y, and at $R_j = 1$ the representation of the species gets value $w_1 + w_2$. Thus, by giving different values to parameters you have practically an endless number of options for the shape of the benefit function. Note however, that the definition of how marginal value is calculated does not change from that of additive benefit function (section 2.3.2.). Also with this cell removal rule species occurrences are considered as additive and the cell that has the lowest marginal value summed across all species will be removed next.

Some shapes that generalized benefit function can assume are listed in the following table:



	w_1	w_2	T_j	x	Y
i) Linear	w_j	0	1.0	1.0	NA, dummy=1.0
ii) Power function (=ABF)	w_j	0	1.0	<1 or >1	NA, dummy=1.0
iii) Mild sigmoid	w_j	same order as w_j	at inclination point	>1	<1, e.g. 1/x
iv) Steep sigmoid – step imitation	w_j	same order as w_j	at step	>>1	<<1, e.g. 1/x
v) Ramp	w_j	0	at step	1.0	NA, dummy=1.0
vi) Ramp, with linear over-representation	w_j	<< w_j	at step	1.0	1

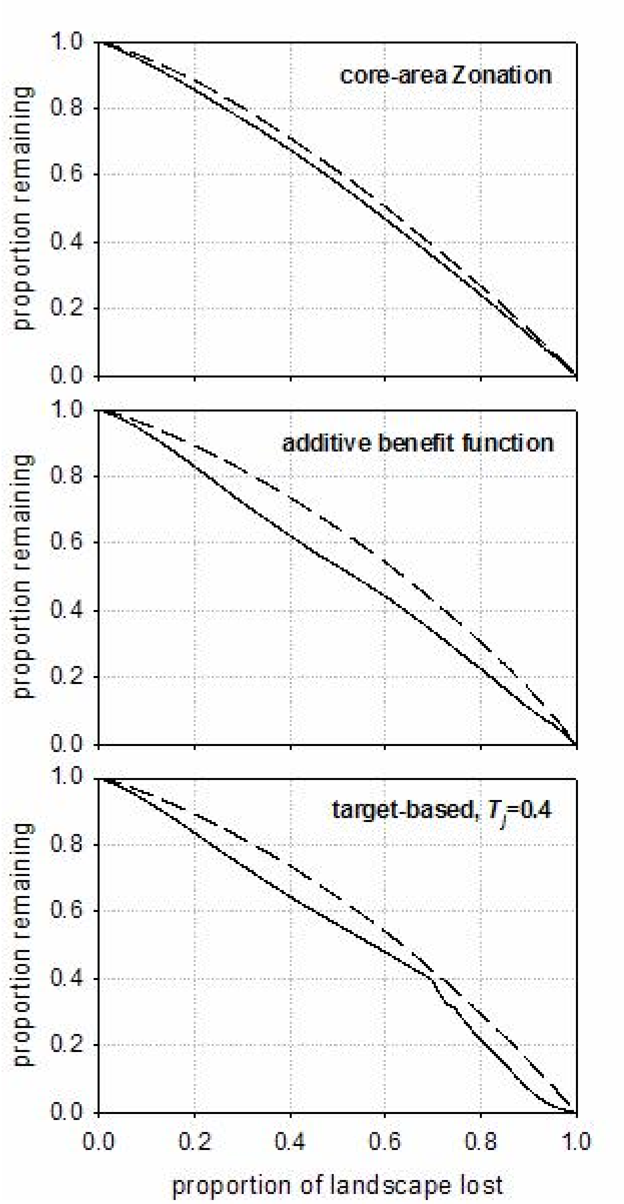
The parameter definitions are suggestive, and the exact shape of the function can easiest be determined by plotting it. To use generalized benefit function as a cell removal rule the parameters of the function need to be given in the biodiversity feature list file (section 3.3.2.2.).

2.3.5 General differences between cell removal rules

It is important to realize that there may be significant differences between different cell removal rule solutions and that the most preferable solution method depends on the goals of planning. Thus, different cell removal rules may be conceptually better suited for different situations.

- **Core-area Zonation** is most appropriate when there is (i) a definite set of species all of which are to be protected - tradeoffs between species are discouraged, (ii) the hierarchy of solutions and easy weighting of species is desired and (iii) importance is given to core-areas - locations with highest suitability for a given species (in terms of abundance or high probability of occurrence); occurrences in cells are not additive meaning that twenty locations with p=0.05 is not the same as one location with p=1.0.
- **The additive benefit function** formulation may be more appropriate when (i) the species are essentially surrogates or samples from a larger regional species pool, and tradeoffs between species are fully allowed, and (ii) the hierarchy of solutions and easy weighting of species is desired.
- **The targeting formulation** is most appropriate when (i) it is accurately known what proportion of the landscape can be assigned for conservation and there is no need for a hierarchy of solutions, (ii) there is a definite set of species all of which are to be protected, (iii) occurrences are additive, and (iv) easy weighting of species is not needed. In target-based planning species weighting is essentially done by giving species different targets.

The figure below (from Moilanen 2007) illustrates some general differences between the core-area Zonation, the additive benefit function formulation and the targeting benefit function. Here, the lines show how large proportion of species distributions is remaining in the landscape as cells are progressively removed. Overall, the additive benefit function has the highest average proportion over all species retained (dashed line), but it simultaneously has the smallest minimum proportion retained (solid line), because it favors species-rich areas over those areas that might be significant for the existence of one or few species, but that otherwise are species-poor. Although the core-area Zonation gives a relatively low average proportion of species distributions, it has a high minimum proportion, because it retains the most significant areas of species (the "core areas") till the end, even though these areas might be unsuitable for all the other species. The targeting benefit function does well in terms of finding the highest level of cell removal without having any species-specific targets violated. However, once representation targets are violated, it performs relatively poorly in terms of the minimum fraction over species retained. The problem with the targeting benefit function is that it is aimed at good performance at one particular set of targets, but the hierarchy of solutions is missing in the sense that good overall performance at other levels of cell removal, especially at a level where targets have been violated, cannot be guaranteed.



There are also differences between the cell removal rules on how much area they require for achieving a set conservation target. To get a given minimum fraction across species core-area Zonation requires more cells than the benefit function variants. This is because benefit function variants take occurrences as additive whereas core-area Zonation prefers the locations with very highest occurrence levels. However, if one investigates the number of cells needed to get the target distribution for an individual species, then core-area Zonation may require fewer cells because it prefers the higher-quality (density) cells. Thus, benefit function variants generate landscapes with many species occurring simultaneously at potentially low occurrence levels and with high overlap between species. Core-area Zonation produces solutions with species occurring at higher densities, but with less overlap between species.

All these differences can logically be expected to occur with any data set, with the magnitude of differences depending on the nestedness of species distributions. Differences would be largest when there are both (i) substantial regional differences in species richness and occurrence levels and (ii) a generally low overlap between species distributions. In this case, the core-area Zonation could catch cores of species occurring in species-poor areas, whereas the additive benefit function would concentrate the solution more towards species-rich locations, where cells have high aggregate value over species.

Core-area Zonation and presence-absence data

When species data is in binary presence-absence form, all cells where a species is present receive an equal value (of 1). In such data, there apparently are no core-areas particularly important for the species, and it might seem pointless to use core-area Zonation as the cell removal rule. This is not the case. Firstly, any additional analyses such as aggregation methods or the uncertainty analysis will cause the value of the cell to be calculated based on other features (e.g. connectivity of the cell) in addition to the species data. Differences between areas where the species is present do therefore emerge. Secondly, because core-area Zonation does not treat species occurrences as additive, but tries to retain occurrences of individual species in the landscape as long as possible, there will still be significant difference in the cell removal process between core-area Zonation and additive benefit functions. We illustrate this with an example. Let us assume we have a landscape where 7 different species occur. Six of these species have overlapping distributions and one (denoted here as species A) has a distribution isolated from the other species. Because benefit functions take species occurrences as additive, the cells in sites where distributions of several species overlap receive a higher value than the cells where only one species occurs, as is the case with species A. Thus, in the cell removal process the additive benefit function would always favor cells with multiple species over the cells of species A, which would lead to unequal preservation of species (in other words species A would lose its distribution much more quickly than the other species). In contrast, Core-area Zonation would retain all species distributions equally, meaning that species A would lose its distribution at the same pace as do the other six species. This conclusion stays the same even when using presence-absence data.

2.3.6 Random removal

The fifth cell removal rule in Zonation is random removal. This option might be useful for quality control, i.e. if you are exploring different methods for cell prioritization and want to find out what would be the baseline representation level that you would get regardless of how efficient methods you are using. Random removal removes cells, as the name implies, in a random order, with no consideration of conservation value of the cells. Here, the assumption with respect to analysis outcome is that the average representation curve should be a straight line going from one to zero when the fraction of the landscape goes, likewise, down from one to zero.

2.4 Inducing reserve network aggregation

Fragmentation is an undesirable characteristic in reserve design as it has been concluded in many studies that species persist poorly in small and isolated patches. Also, implementing a fragmented reserve network may be awkward and expensive. In this section we introduce three different aggregation methods that can be used when running analyses with the Zonation program. These methods produce relatively more compact solutions. Note, however, that aggregation always involves trade-offs. There is usually an *apparent* biological cost in more aggregated solutions because in many cases it is necessary to include lower-quality habitats into the reserve network in order to increase connectivity. In reality this apparent loss is more than offset by benefits of having a well-connected area. Thus, it is recommended to use aggregation methods in reserve planning as the cost of loosing a minor amount biologically valuable areas is usually low compared to the benefits of high connectivity. For more information on true and apparent costs related to aggregation, see Moilanen and Wintle (2006) and (2007).

There are some distinct differences between the aggregation methods in Zonation, and choosing the right one depends on conservation targets and computational issues.

- **Boundary Length Penalty** (BLP) has been the most commonly used way to introduce aggregation to reserve planning. However, it is important to understand that BLP is a general, non-species-specific aggregation method which does not asses the actual effects of fragmentation on species. Rather the method only uses a penalty on a structural characteristic of the reserve network (boundary length) to produce more compact reserve network solution. The method is computationally quick and effective, but might not be biologically most realistic.

- **Distribution Smoothing** is a species-specific aggregation method which retains areas that are well connected to others, thus resulting a more compact solution. The connectivity of sites is determined with a smoothing kernel, which means that the value of a cell is "smoothed" to the surrounding area. Another way of looking at distribution smoothing is, that it does a two-dimensional habitat density calculation, identifying areas of high habitat quality and density. Consequently, cells that have many occupied cells around them receive a higher value than the isolated ones. The widths of the smoothing kernel are species-specific, implicitly expressing the species dispersal capability or scale of landscape use. This aggregation method is computationally very quick. However, it assumes that fragmentation (low connectivity) is generally bad for all species and it always favors uniform areas over patchy ones.

- **Boundary Quality Penalty** (BQP) is biologically the most realistic aggregation method included in Zonation. This method describes how the local value of a site for a species is influenced by the loss of surrounding habitat. The change in local value is based on species-specific responses to neighborhood habitat loss, thus local value may also increase if the site includes species that benefit from fragmentation. The downside of this method is the required computation time, which is much higher compared to the other two aggregation methods. This is because each cell removal influences the habitat value in all remaining neighborhood cells, which needs to be accounted for in the cell removal process.

- **Directed connectivity** (neighborhood Quality Penalty; NQP) is a generalization of boundary quality penalty, where the connectivity between sites is strictly directed, such as in riverine systems. This option also demands the use of planning units, groups of cells, which are removed as a whole instead of singular cells during the landscape ranking process. (In freshwater planning these units would correspond to catchments.) The value of a focal planning unit is influenced by the removal of other planning units "upstream" or "downstream" of the focal unit. Following the philosophy of BQP, also here the change in local value is based on species-specific responses to nearby habitat loss. However, the computation times are relative to the size of the planning units: Smaller planning unit size means longer computation times and vice versa.

- **Matrix connectivity** is a connectivity calculation, where the local value for the focal feature is multiplied by connectivity to multiple other features simultaneously. For example, this can be useful when different but similar habitat types contribute to each other's connectivity. For example, mixed spruce forest next to mixed pine forest is almost the same as homogeneous habitat for many species. This feature can also be used to induce heterogeneity, when it is desirable to have a mixture of certain habitat types rather than a homogeneous landscape.

Simultaneous use of multiple aggregation methods will cause difficulties for interpreting your results. We therefore recommend that you are careful with this. However, there are no technical reasons why smoothing, the BLP, the BQP and matrix connectivity could not be used in the same analysis.

2.4.1 Boundary Length Penalty (BLP)

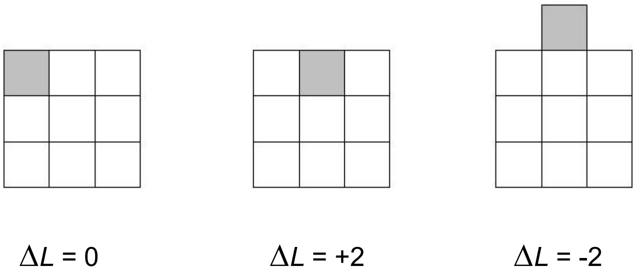
This section is mainly based on Moilanen and Wintle (2007).

Boundary length penalty has been the most common way of adding aggregation into a reserve network. This method is qualitative in the sense that the estimated conservation value of individual cells, or consequently the conservation value of the entire reserve network, is not influenced by the degree of fragmentation, but rather, aggregation is induced via a penalty given for the boundary length of the reserve. BLP does not include species- or habitat-specific components.

When using the boundary length penalty, the hierarchy of cell removal is based on both species occurrence levels in cells and the increase/decrease of boundary length that results from the removal of a cell. The boundary length penalty can in the context of core-area Zonation be formulated as:

$$\delta_i = \max_j \frac{Q_{ij}(S)w_j}{c_i} + \beta \cdot \Delta(BL / A),$$

where $\Delta(BL/A)$ is the change in boundary length/area -ratio of the reserve network following removal of cell i, and β is a constant defining the strength of the boundary length penalty. If cell removal decreases boundary length, $\Delta(BL/A)$ receives a negative value and the value of δ_i for cell i decreases, indicating that it is relatively advantageous to remove the cell because removing it reduces fragmentation.



Above is a picture showing how different cell removal options would influence the boundary length. The boundary length is calculated in the terms of cell edges. Removing the gray cell in the first example results in no changes in boundary length for as two edges are removed while another two are gained. In the second example the cell removal leads to the loss of one edge, but also to the gain of three new edges. Thus as a result the total change in boundary length is +2, and so on. To get from ΔL *to* $\Delta(BL/A)$ one needs to account for both the change in boundary length and the decrease of the reserve area by one.

Note that BLP is different from both distribution smoothing and the BQP. First, the BLP is not a species-specific way of handling reserve connectivity. It simply uses a penalty term that devalues reserve structures with lots of edge. This is completely qualitative; there is no species-specific parameter or response. Like distribution smoothing, the BLP may be expected to perform poorly for species that happily occur in fragmented habitats. This is because the BLP qualitatively favors structurally connected areas, and it will therefore have a tendency to remove small habitat fragments from the solution irrespective of whether some species can actually persist in them or not.

Instructions to using BLP in Zonation can be found in section 5.1.2.

Literature

Boundary length penalty is described by
Moilanen, A., and Wintle, B. A. 2007. The boundary-quality penalty: a quantitative method for approximating species responses to fragmentation in reserve selection. *Conservation Biology*, 21: 355-364.

2.4.2 Distribution smoothing

This section is mainly based on Moilanen *et al.* (2005) and Moilanen and Wintle (2006).

With distribution smoothing, planning is based on a connectivity surface computed from the original species distributions that have been input into Zonation. The calculation applied to each species distribution is identical to the calculation of a metapopulation-dynamic connectivity measure, where the connectivity value is directly proportional to the number of migrants expected at a given location in the landscape. Technically, the computation is a two-dimensional kernel smoothing using a species-specific parameter (width of the smoothing kernel).

For practical purposes, distribution smoothing identifies areas that have on average high occupancy levels for species. The smoothing very effectively identifies important semi-continuous regions where the species has overall high levels of occurrence, although not necessarily in every grid cell. In contrast, relatively scattered occurrences in fragmented habitat lose value. As the distribution of the species in the landscape becomes smoother, populations in fragmented areas end up with less value than continuous areas with the same average probability of occurrence.

Note that distribution smoothing should be used with care if the data includes a species that lives happily as a metapopulation in a fragmented habitat - smoothing should be narrow for this species at least if the habitat matrix is taken as partially suitable for the species. Distribution smoothing is a convenient technique to apply because it can be run as a relatively fast preprocessing step before going on to the Zonation analysis itself. The appropriate level of smoothing for a given species would be determined based on a conception of the typical dispersal distances for that species, or from information concerning home range sizes for the species.

When using smoothing, the value for species *j* in a focal cell *i* is

$$O_{ij}' = \sum_{x} \sum_{y} \exp(-\alpha_j d(x-u, y-r)) O_{ij} ,$$

where O_{ij} is the original occurrence level of species *j* at cell *i*. Cell *i* is located in (*u*,*r*) and *d*(*x*-*u*, *y*-*r*) is the distance between locations (*x*,*y*) and (*u*,*r*). The summation is over the landscape grid and α_j is the parameter of the dispersal kernel for species *j*. This is a two-dimension kernel smoothing using a radially symmetric negative exponential (dispersal) kernel.

Instructions to using distribution smoothing in Zonation are in section 5.1.3.

Literature

For details about distribution smoothing, see

Moilanen, A., Franco, A. M. A., Early, R., Fox, R., Wintle, B., and Thomas, C.D. 2005. Prioritising multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society of London, Series B, Biological Sciences* 272: 1885-1891.

Moilanen, A. and Wintle, B.A. 2006. Uncertainty analysis favors selection of spatially aggregated reserve structures. *Biological Conservation,* 129: 427-434.

2.4.3 Boundary Quality Penalty (BQP)

This section is mainly based on Moilanen and Wintle (2007).

The boundary quality penalty is a quantitative species-specific way of inducing aggregation into Zonation solutions. It can be seen as a way of approximating nonlinear effects of connectivity that may be present in habitat models. The rationale behind the BQP goes as follows.

There are very many different statistical species distribution modeling techniques (a.k.a. habitat models, resource selection functions etc.). Typically, in such models the abundance of a species at a location is influenced not only by local habitat quality, but also by habitat in the neighborhood of the location. Such a neighborhood influence essentially states that the species is somehow dependent on connectivity or edge effects or both.

Now, ideally, reserve selection would be directly based on nonlinear habitat models with neighborhood effects. However, this is not realistically possible, because it would make reserve selection computationally very, very slow. Also, implementing dozens of different habitat modeling techniques efficiently into reserve selection software would be an enormous task. Herein enters the BQP.

The BQP is a mechanism for approximating the aggregate response of a species to edge effects and metapopulation size and connectivity. It can be seen as a way of exploiting the connectivity response that is present for a species in a habitat model. Essentially, one uses the habitat model for two things. First, one predicts an abundance or probability of occurrence into every cell in the landscape – giving the standard input layer for one species. Additionally, one analyses the habitat model to find out what kind of an aggregate response to habitat loss and fragmentation does the species have. And this response is transferred into Zonation as a standardized curve, which mediates the boundary quality effect in Zonation. [Note: we emphasize that the BQP responses for species can also be generated based on expert opinion; they do not always need to be statistically fitted.].

Different species can have different responses to fragmentation and habitat loss, which are entered into Zonation as two BQP components: (i) a species-specific radius and (ii) a response curve. The radius is species-specific and defines the distance from which habitat loss has an effect on a species in a cell. The effect can be very localized (e.g. if the species is only sensitive to edge effects very close to the focal cell), or, the effects can extend over a long distance. This could be the case with a timid larger animal that avoids human proximity - habitat loss could influence such a species from a long distance from the cell where it actually occurs. Note that inside the radius only the loss of those cells that have data on the particular species (in other words cells that are not marked as missing data) can influence the value of the focal cell via connectivity.

The second component, the response curve, specifies the kind of effect neighborhood habitat loss has on the species. First, there could be an absence of effect, which would be modeled by a flat line (no effect). Then, the species could suffer variable degrees of loss in local population density if habitat is lost in the neighborhood. For example, you could have a relatively insensitive species, which loses half of the population density when the focal cell has lost all it neighbors from inside the species-specific buffer. Then again, the species could be very sensitive to neighborhood fragmentation – all local value could be lost when only half of the neighboring cells have been lost. And finally, a species could even favor fragmentation, which would be modeled by a response curve that goes above 1 at some levels of habitat loss.

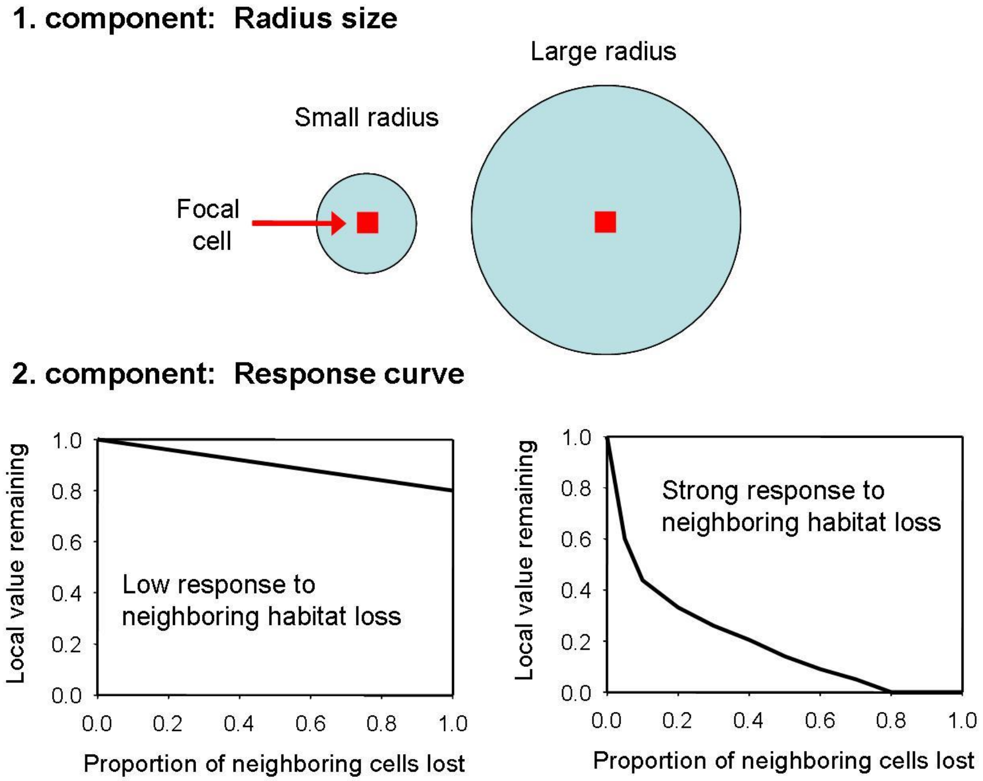

1. component: Radius size
Large radius
Small radius
Focal cell
2. component: Response curve
Local value remaining
1.0
0.8
0.6
0.4
0.2
0.0
Low response to neighboring habitat loss
0.0
0.2
0.4
0.6
0.8
1.0
Proportion of neighboring cells lost
Strong response to neighboring habitat loss

Figure clarifying the two components used when specifying the BQP for a species.

For the exact way of analyzing a habitat model to get the species-specific response, see Moilanen and Wintle (2007). Note also that the BQP curves need not be derived from habitat models. One could also use expert opinion to guesstimate the response of the species to neighborhood habitat loss. The hypothetical curve would then be entered into Zonation – it makes no difference for the Zonation process how the BQP-curves were obtained.

Note that there are major differences between distribution smoothing and the BQP even though both induce aggregation in a species-specific manner. First, from a practical point of view, the BQP is much slower to run as the effect of removing a cell is not only local but extends over the neighborhood area which needs to be accounted for in computations. Second, the BQP is biologically better justified. The BQP definitions can be based directly on species responses in statistical habitat models. The difference between the methods would be most strikingly visible in fragmented areas. Distribution smoothing perceives the value of fragmented areas as relatively low. In comparison, the BQP could recognize a species that happily lives as a metapopulation in a fragmented environment – the response for that species would be such that it is recognized that the species can have high-value habitats in fragmented areas.

In the implementation of the BQP into Zonation, the value of a cell that is removed (marginal loss of conservation value) is now divided into two components (i) local value, which is as before, and (ii) loss of conservation value in the neighborhood of the focal cell as modeled via the BQP specification. Thus with BQP the effect of cell removal is not only the loss of the value in the cell itself but also a (potentially species-specific) reduction in quality in the neighborhood cells:

$$\delta_i = \max_j \frac{w_j}{c_i} \left\{ H_j\left(\frac{n_{ij}}{n_{ij}^{*}}\right) Q_{ij}(S) + \sum_{k \in N_j(i)} Q_{kj}(S) \left[H_j\left(\frac{n_{kj}}{n_{kj}^{*}}\right) - H_j\left(\frac{n_{kj}-1}{n_{kj}^{*}}\right) \right] \right\},$$

where $N_j(i)$ indicates the cells (containing data for that species) within the species-specific radius of cell i for species j. Denoting by h the fraction of original neighbors that have been lost from within the species-specific buffer around the site, $H_j(h)$ is the proportion of the original value of cell k remaining for species j when the focal cell has fraction h of its neighbors remaining. The fraction of cells remaining is simply $h = n_{kj} / n_{kj}^*$, where n_{kj} is the number of neighbors remaining for cell k within the buffer radius of species j, and n_{kj}^* is the original number of neighbors. The loss term in the curly brackets is divided into two, local loss and loss in the neighborhood of the focal cell. Local loss is the fraction remaining of the original value of the focal cell i - if many of its neighbors have already been lost, the value of $Q_{ij}(S)$ has been reduced. Loss in the neighborhood is mediated via the loss of one cell from the number of neighbors, which goes down from n_{kj} to n_{kj} - 1. (Note that the formula above is employed as it is only for the core-area Zonation. For additive benefit function and target-based planning the formula includes few trivial differences (see Moilanen 2007), but the concept behind BQP is the same in all cases.)

The size of the neighborhood of a cell and the effects of habitat loss are defined separately for each species according to habitat models, which themselves mediate the boundary quality penalty. Because BQP ranks the cells based on the responses of species to fragmentation, also species that actually benefit from the loss of surrounding habitats will be equally protected (compare distribution smoothing).

Instructions to using BQP in Zonation can be found in section 5.1.4.

Literature

Boundary quality penalty is described by

Moilanen, A., and Wintle, B. A. 2007. The boundary-quality penalty: a quantitative method for approximating species responses to fragmentation in reserve selection. *Conservation Biology*, 21: 355-364.

2.4.4 Directed connectivity (NQP)

This section is mainly based on Moilanen, Leathwick, and Elith (2008).

The directed freshwater connectivity measure is a generalization of the BQP technique, in which the concept of neighborhood is generalized – hence the name Neighborhood Quality Penalty (NQP). Instead of using a circular neighborhood, the NQP is defined using a tree-hierarchy of linked planning units. A focal area (planning unit) is influenced by negative action (habitat loss) potentially both “downstream” and “upstream” from the focal location, depending on the requirements of the species.

The NQP technique was originally developed for freshwater planning in riverine systems.

Note however, that the technique could be suitable for quite different situations as well: the NQP method is based on a bidirectional linking of planning units. These linkages could correspond to hydrological flow. But, they could also correspond to other kinds of biological linkages, including (i) other spatially continuous connecting landscape elements, such as hedge rows, or (ii) spatially discontinuous, but functionally linked planning units, such as areas on migration routes of birds, or (iii) they could approximate connectivity at marine areas where very strong flows generate a situation analogous to a river system. In the end, an appropriate aggregation of cells to planning units and suitable linkages and loss functions allow modeling of relatively variable situations.

Following closely Moilanen *et al*. (2008), the present version of the NQP technique is technically specified by the following modification of the marginal loss value used in the cell removal rule:

$$
\begin{aligned}
\delta_{ij} &= \delta_{ij}^{local} + \delta_{ij}^{neighborhood} = \delta_{ij}^{local} + \delta_{ij}^{upstream} + \delta^{downstream} \\
&= p_{ij} h_j^{up}\left(\frac{r_{ij}^{up}}{o_{ij}^{up}}\right) h_j^{down}\left(\frac{r_{ij}^{down}}{o_{ij}^{down}}\right) \\
&+ \sum_{k \in N_{ij}^{up}} p_{kj}\left[h_j^{up}\left(\frac{r_{kj}^{up}}{o_{kj}^{up}}\right) h_j^{down}\left(\frac{r_{kj}^{down}}{o_{kj}^{down}}\right) - h_j^{up}\left(\frac{r_{kj}^{up}}{o_{kj}^{up}}\right) h_j^{down}\left(\frac{r_{kj}^{down} - A_{ij}}{o_{kj}^{down}}\right)\right] \\
&+ \sum_{k \in N_{ij}^{down}} p_{kj}\left[h_j^{up}\left(\frac{r_{kj}^{up}}{o_{kj}^{up}}\right) h_j^{down}\left(\frac{r_{kj}^{down}}{o_{kj}^{down}}\right) - h_j^{up}\left(\frac{r_{kj}^{up} - A_{ij}}{o_{kj}^{up}}\right) h_j^{down}\left(\frac{r_{kj}^{down}}{o_{kj}^{down}}\right)\right].
\end{aligned}
$$

(1)

in which p_{ij} is the occurrence level of species *j* in cell *i*. Equation (1) describes the fraction lost from original distribution of species *j* following the removal of site *i*. The loss consists of three components, local loss, loss upriver, and loss downriver. The assumption is that everything remaining locally is lost if a cell is removed, and that loss accrued upriver and/or downriver will depend on the size of unit *i* for species *j*, A_{ij} – loss of a larger unit implies greater influence on connectivity nearby. The influence of connectivity on occurrence levels is mediated via functions h_j (upriver and downriver), which are response functions like those in BQP with the x-axis reversed. When the full landscape remains, and nothing has been lost, h_j (1)=1. Importantly, when calculating marginal loss, the equation accounts for degradation that already has occurred. This implies that if no local value remains due to past neighborhood loss, further loss of connectivity has no local influence on the species.

Connectivity in Eq (1) is modeled separately upriver and downriver. Quantities r_{ij} and o_{ij} (up and down) are the remaining and original connectivities of unit *i* for species *j* both upriver and downriver, respectively. Loss of planning unit *i* influences the downwards connectivity of sites upriver from it, which is the component of the equation having the summation across neighbors *k* upriver from focal site *i*, $k \in N_{ij}^{up}$. Similarly, loss of unit *i* influences the upwards connectivity of units downriver from it. Note that the present version of Zonation uses species and unit specific predictions of occurrence p_{ij} and species-specific connectivity responses h_j(). However, connectivity up and down is based directly on the numbers of grid cells in planning units in a non-species-specific manner, meaning that r_{ij}, o_{ij} and A_{ij} are taken as the same for all species, which assumption might be relaxed in a later version of Zonation.

Equation (1) is simply the fraction of distribution lost for one species, which does not account for how lost representation is translated to lost in conservation value. When deciding which cell can be removed with smallest loss of conservation value, δ_{ij} is aggregated across species according to the cell removal rule, which now is for core-area Zonation (section 2.3.1.)

$$\delta_i = \frac{1}{c_i} \max_j \frac{w_j \delta_{ij}}{q_j}, \tag{2a}$$

and for the additive benefit function Zonation (section 2.3.2.)

$$\delta_i = \frac{1}{c_i} w_j \sum_j \left[V_j(q_j) - V_j(q_j - \delta_{ij}) \right], \tag{2b}$$

in which w_j is weight of the species j, c_i is cost (or area) of planning unit i, and q_j is the fraction of distribution of species j remaining before removal of cell i, and $V_j()$ is the function translating increasing representation into increasing conservation value (see sections 2.2.1. and 2.2.2.; Arponen *et al*. (2005); Moilanen (2007)).

Instructions for using directed connectivity in Zonation can be found in section 5.1.5.

Literature

For more information about the method, see

Moilanen, A., Leathwick, J. and J. Elith. 2008. A method for freshwater conservation prioritization. *Freshwater Biology*, 53: 577-592. *Freshwater Biology*, 53: 577-592.

2.4.5 Matrix connectivity

This section follows Lehtomäki et al. 2009.

A new way of accounting for connectivity in Zonation. The rationale behind this technique is that multiple features can facilitate connectivity for each other. For example, Lehtomäki et al. (2009) applied matrix connectivity to account for the extent to which different types of forest habitats enhance connectivity of each other. One may divide the landscape to different forest types, but, while a spruce forest may be different from a pine forest, they would still influence each others connectivity positively, compared to the non-forest habitats such as agricultural fields or water bodies.

Matrix connectivity is essentially a multi-feature extension of ecological interactions, where the conservation value of a cell for a species is affected by its connectivity to cells where interacting features (do or do not) occur within the dispersal kernel of the species (see section 2.5. for a detailed description of ecological interactions and section 2.3.2. for distribution smoothing). The value p' of cell i for species or other biodiversity feature *k* is calculated as

$$p'_{ik}=p_{ik}C_{ik} \qquad (1),$$

where p_{ik} is the occurrence level or other quality measure for biodiversity feature *k* in cell *i* and C_{ik} is the multi-feature connectivity of cell *i* from the perspective of feature *i*. The Zonation algorithm first normalizes p_{ik} to the fraction of the full distribution of feature *k* in cell *i*. The important part here is the way C_{ik} is defined:

$$C_{ik} = \sum_{n=1}^{F} \{S_{nk} \sum_{j=1}^{J} p_{jn} \exp [\, \alpha_k \, d(i,j)]\} \qquad (2),$$

where *F* is the number of all features, *J* is the total number of cells, p_{jn} denotes the occurrence level of feature *n* in cell *j*, *d(i,j)* is the geographical distance between cells *i* and *j*, and α_k is the parameter giving the spatial scale for feature *k*. S_{nk} is a coefficient specifying how much feature *n* contributes to the connectivity of feature *k*. S_{nk} values are read in to Zonation from a connectivity matrix (section 3.3.3.4.).

Instructions for running an analysis with matrix connectivity are in section 5.1.6.

Literature

Matrix connectivity has been described by

Lehtomäki, J., Tomppo, E., Kuokkanen, P. Hanski, I., and A. Moilanen. 2009. Planning of forest conservation areas using high-resolution GIS data and software for spatial conservation prioritization. Forest Ecology and Management, 258: 2439-2449.

2.4.6 Edge adjustment in connectivity

This section is mainly based on Arponen, A., Lehtomäki, J., Leppänen, J., Tomppo, E. and Moilanen, A. Submitted manuscript. Analysis resolution and connectivity in large-scale spatial conservation prioritization.

This is an adjustment to feature-specific matrix connectivity values of cells at edges of habitat. Connectivity decreases towards edges, but in some cases it may not make sense to consider edges to have lower priority. There are two typical cases where such an adjustment may be used:

(i) Cells on the other side of the national border could be marked as non-harmful base habitat even if exact biodiversity data is used. Then, it is effectively assumed that habitat outside the border will influence connectivity as habitat inside the border.

(ii) At the edges of water bodies. For example, thinking of connectivity of a forest. The connectivity of forest will necessarily be reduced at the edge of a large lake. But, there may be many cases where such an edge-effect is not desirable. If not, mark both forest and water as suitable base habitat in the connectivity edge effect fix file. Or, this option could be relevant at the border of a forest and a marshland (some species will not perceive the marshland as bad for connectivity).

In practice, specific areas of non-habitat are ignored in matrix connectivity calculations. Connectivity values are corrected by the fraction of landscape around the focal cell from which connectivity is aggregated. If the neighborhood of the focal cell includes what can be considered harmless non-habitat (e.g. habitat continues across country border unchanged), then the connectivity value is corrected to $c'=c/(1-f)$, where c is original connectivity and f is the fraction of harmless non-habitat. For example, if the connectivity of the cell is 2.0, which has been aggregated from a neighborhood that is only 1/3 within the country, then if cells outside the country have been marked as non-harmful base-habitat, the value of connectivity becomes 2.0/(1/3)=6.0.

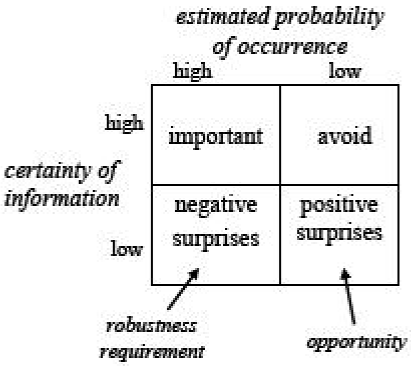
In addition, it is possible to correct for the proportion of habitat within each cell. When 50% of a cell is located outside country border, or when half of it is covered by water, its biodiversity value cannot be as high as it is for cells covered entirely by habitat. This can be done through the use of a cost layer, where cost is the proportion of habitat per cell. In practice, the size of the occurrence of a feature in a cell is divided by the proportion of habitat in the cell, transforming absolute habitat amount to habitat density. For example, if a cell's biodiversity value as such was 5 but only 50% of it actually contains habitat and the rest is covered by a lake, then this correction increases the cell's value to 10. The other way to think of it is that the cost of protecting the 50% of the cell that actually contain habitat should be 50% of land cost per cell in the region.

It is up to the user to decide when these edge related adjustments are desirable and when not. The latter correction considers "quality" of cells, but not actual land costs. Land cost, when known, should be included as well, in practice as a product with habitat proportion in a single cost layer.

2.5 Uncertainty analysis

A common problem with conservation planning is the uncertainty of planning inputs. Mostly these uncertainties are due to lack of data - we simply do not have a comprehensive database with accurate information of the distribution of every species. Uncertainty can also arise, for example, from outdated or false observations, the use of predicted data (e.g. distribution models), or from any future factors such as the potential for anthropogenic land-use changes, or climate change.

Taking into account both biological value and uncertainty creates a prospect of four scenarios. (1) Areas with high conservation value and high certainty of that information would be important for conservation. (2) Areas with low conservation value and high certainty (car parks, etc.) would ordinarily rank low among conservation priorities. (3) Areas with high estimated conservation value, but low certainty have potential for producing negative surprises for conservation. (4) Finally, areas with low conservation value and also low certainty have potential for producing positive surprises.



The goal of uncertainty analysis in reserve selection is to implement and evaluate trade-offs between biological quality and the certainty of that information. Ideally, one would identify a reserve network that guarantees high biological quality despite some uncertainty in input data. Uncertainty analysis could also be used for evaluating the opportunities arising from uncertainty, that is, potential for positive surprises.

Here we introduce three methods of uncertainty analysis that can be used in the Zonation context. The first method is called distribution discounting, which enables the ranking of the landscape using species distribution data that includes uncertainties. The second is called opportunity analysis, and it differs from the previous by giving high value to low uncertainty, with the aim of maximizing positive surprises. Our third method is for testing reserve network structures to see how robust they are against uncertain negative influences of habitat loss and fragmentation.

Literature

For more information about the aims and methods of uncertainty analyses in reserve selection, see:

Moilanen, A., Runge, M. C., Elith, J., Tyre, A., Carmel, Y., Fegraus, E., Wintle, B., Burgman, M. and Y. Ben-Haim. 2006. Planning for robust reserve networks using

uncertainty analysis. *Ecological Modelling* 199: 115-124.

Moilanen, A. and Wintle, B.A. 2006. Uncertainty analysis favors selection of spatially aggregated reserve structures. *Biological Conservation* 129: 427-434.

Moilanen, A., Wintle, B.A., Elith, J. and Burgman, M. 2006. Uncertainty analysis for regional-scale reserve selection. *Conservation Biology* 20: 1688-1697

2.5.1 Uncertainty in species distributions, distribution discounting

This section is mainly based on Moilanen *et al.* (2006).

Distribution discounting is a method for including uncertainty analysis into the conservation prioritization done in Zonation. This method helps you find the most robust solutions, those that most likely achieve a conservation goal given a level of uncertainty in species distributions. This analysis utilizes both the estimated biological value (probability of occurrence) of a species in a cell, and the certainty of that information.

Looking for robust reserve networks

In the framework of uncertainty analysis, one goal for reserve selection would be to find those network candidates that would achieve the given conservation targets despite uncertainty in input data. Thus, cells need to be ranked so that the highest priority is given to cells that have both relatively high conservation value and high certainty of information. In Zonation, uncertainty analysis has been implemented according to a convenient formulation that uses information-gap decision theory (see Ben-Haim 2006).

Conceptually, relevant components of the info-gap theory are

1. The *nominal model*. This is your best set of predictions for species.

2. The *uncertainty model*. This states that even though you do have a nominal estimate, your true probabilities of occurrence are certain to deviate from your nominal model. The uncertainty model specifies a set of bounds that expand around the nominal estimate as a function of an uncertainty parameter, α.

3. A *performance function*. This is a function measuring how well you are doing, for example, what is the proportion of species distributions that would be covered by a given set of areas.

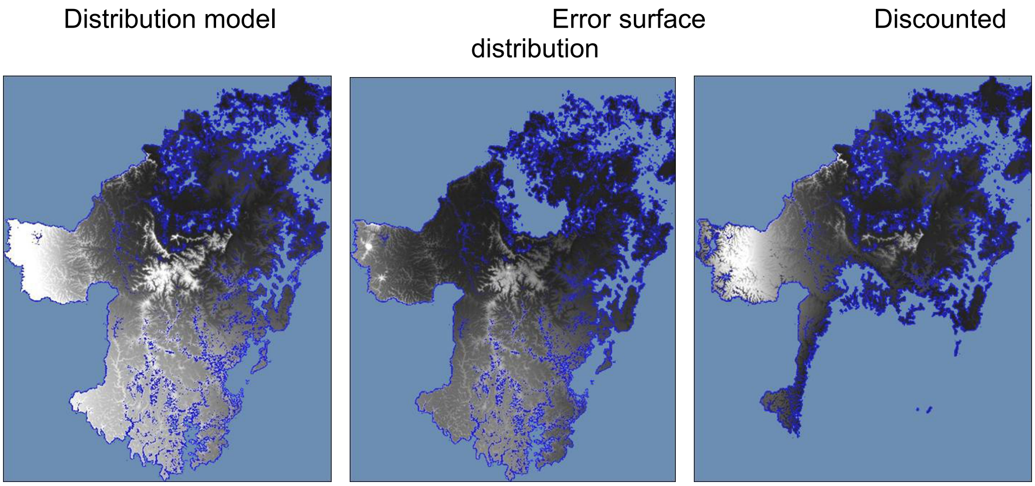
4. *Robustness function*. This function measures "how large can the horizon of uncertainty α be so that conservation goals are still met even if you take the most adverse choice of probabilities from inside the uncertainty bounds". A good reserve candidate is such that it achieves goals while allowing for high uncertainty (α). The *robust optimal reserve candidate* is the one that achieves conservation goals while allowing for highest uncertainty.

At simplest, distribution discounting is implemented as follows

1. Take your nominal estimates (the normal input distributions for species)

2. Develop a respective uncertainty map for each species. The uncertainty layer could, for example, represent the standard deviation (SD) of the nominal estimate or the length of the lower half of the 95% confidence interval. This map could also be based on unstatistical uncertainty, for example, on an expert-based estimation on the likelihood of an area becoming damaged in the near future.

3. Specify, e.g., that the horizon of uncertainty *α* is 0.5 or 1. E.g., if $\alpha = 1$ and your uncertainty model is 1xSD, then you would essentially subtract one SD of the nominal estimates, thus emphasizing locations with relatively certain predictions.



Distribution model Error surface Discounted distribution

Picture demonstrating the concept of distribution discounting. Here the first picture shows a modeled map of species distribution, white areas representing a high probability of occurrence. The second picture displays an error surface, eg. standard deviation of the modeled values, again white color indicating large deviation and therefore high uncertainty. The *α*-value (horizon of uncertainty) can be used to either enhance or diminish the strength of the error surface, e.g. $\alpha = 2$ would double all error values in the map. The third picture is species discounted distribution, where the error surface has been subtracted from the modeled distribution map. This is the map that Zonation finally uses to run the analysis.

An expanded explanation for distribution discounting.

In more detail: The occurrence of species *s* in a cell *c* (here indicated as p^*_{sc}) is by no means certain, but merely the nominal "best guess" probability. Thus the true probability $p_{sc} \in [0,1]$ could be within an interval given by:

$$\left(p^*_{sc} - \alpha w_{sc}\right) \le p_{sc} \le \left(p^*_{sc} + \alpha w_{sc}\right) \qquad (1)$$

where *α* is the horizon of uncertainty and w_{sc} is any error measure related to the accuracy of p^*_{sc} (for species *s* in cell *c*). Thus the true probability p_{sc} could be either higher or lower than the estimate p^*_{sc}, with bounds for p_{sc} determined by *α* and the relative error measure w_{sc}, which could be, for example, related to the accuracy of statistical prediction. The model of Eq. (1) is called a uniform bound model in info-gap terminology.

When using predictions based on logistic regression habitat models, a plausible model for uncertainty is to define the uncertainty interval in logit space, where w_{sc} is the standard error for the linear predictor of a logistic regression:

$$\left|\text{logit}\ (p_{sc}) - \text{logit}\ (p^*_{sc})\right| < \alpha w_{sc} \qquad (2)$$

According to info-gap theory, one should favor reserve structures that achieve given conservation targets even with the most adverse choice of probabilities (in other words, in the worst case scenario). Given the present definitions, the most adverse choice of probabilities occurs when all probabilities are at their lower bounds (this is when the lowest expected number of populations is obtained). Assuming the analysis in logit space:

$$\text{logit}\ (p_{sc}) = \text{logit}\ (p^*_{sc}) - \alpha w_{sc} \qquad (3)$$

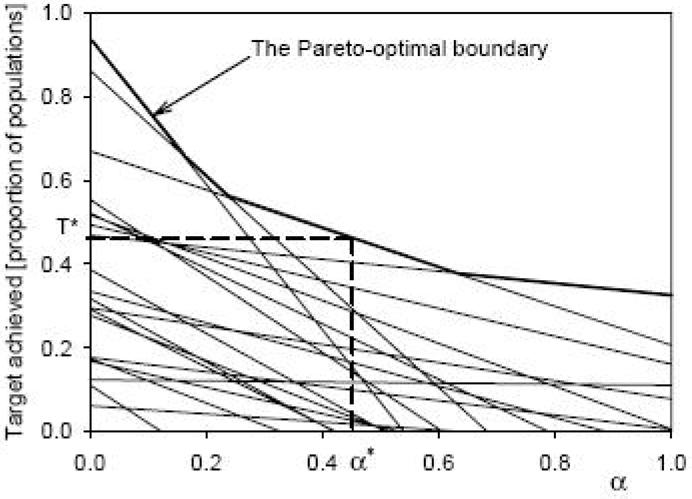
Thus the program calculates the discounted biological value of a cell by reducing (discounting) the value of the logit of probability (p^*_{sc}) by a multiple of the error (αw_{sc}). In the distribution discounting technique, the original estimated occurrence data is simply replaced by the discounted data before proceeding to do the Zonation run. Thus, one Zonation run with discounted data is needed for each value of the horizon of uncertainty, α . Note that Zonation does not care how the nominal estimates and associated error measures are obtained - any statistical method or expert evaluation can be used as a basis for developing those quantities.

But how to determine the horizon of uncertainty, α? As mentioned above, the relative error w_{sc} can be any error measure related to the predicted species distribution. It can, for example, be a statistical error (e.g. the length of the lower half of the 95% confidence interval) or a probability of future anthropogenic threat, or both. The value of α, on the other hand, is unknown and has no correct value. The way forward is to investigate how different reserve network candidates perform under increasing uncertainty. In practise, the way is to try out several levels of α to see what areas are selected. If an area is always selected irrespective of the value of α, then the area is important for sure. If the area is selected with low α but not with high alpha, then the area is selected because of the occurrence of an uncertain biological feature. If the area is not selected with low α but is selected with high alpha, then something of interest occurs in the area with relatively low density but high certainty. Based on such uncertainty analysis, one can split the landscape in three categories of areas: (i) areas that are good for sure, (ii) areas that are never selected and thus have low priority and (iii) areas that are selected with some levels of uncertainty - these areas may need further investigation before a choice of conservation priority for them can be fixed.

Distribution discounting uses the following algorithm for finding robust-optimal reserve designs:

1. Specify robustness requirement α.
2. *Distribution discounting.* Read in species information. For every species and cell, apply Eq. (3) or its analogue, depending on type of data/uncertainty model.
3. Use any reserve selection algorithm (here Zonation) to search over spatial patterns. The robust optimal design **X*** (at the given level of α) is the one achieving the highest possible conservation value.

The advantage of this approach is that the worst-case probability set (Eq. 3) has to be calculated only once (item 2.), and thereafter the contributions of cell to representation levels, p_{sc}, do not change in the reserve selection process. Testing several α-values allows you to outline how different reserve structures behave in increasing uncertainty. Some designs are always bad, some are good according to nominal habitat model predictions, but bad if uncertainty is incorporated into the models. Others have intermediate nominal performance, but have a good robustness to uncertainty. The robust-optimal designs are always at the **Pareto-optimal boundary** with respect to the target, as demonstrated below:



In the figure, each thin line represents one solution (spatial reserve structure). An increasing robustness requirement (α) implies that a decreasing biological value can be achieved reliably. The thick line is the Pareto-optimal boundary representing solutions that are optimal in the sense that increased biological value can only be obtained with the cost of lowered robustness, and *vice versa*. When doing reserve selection on a large grid, there is a huge number of potential reserve structures, but only one or few of them would correspond to the Pareto-optimal boundary at any given α-level and resource (fraction of landscape). Any solution not at the Pareto-optimal boundary is inferior in the sense that another solution exists with either higher biological value or higher robustness or both. The distribution discounting technique used inside Zonation automatically identifies the robust-optimal nested Zonation set of solutions for the given level of α.

Instructions for using distribution discounting in Zonation can be found in section 5.1.8.

Literature

For more information about distribution discounting, see

Moilanen, A., Runge, M. C., Elith, J., Tyre, A., Carmel, Y., Fegraus, E., Wintle, B., Burgman, M. and Y. Ben-Haim. 2006a. Planning for robust reserve networks using uncertainty analysis. *Ecological Modelling,* 199 (1): 115-124.

For more information about information-gap theory in general, see

Ben-Haim, Y. 2006. *Info-gap decision theory: Decisions under severe uncertainty*, 2nd edition. Elsevier Academic Press, London.

2.5.2 Positive uncertainty in species distributions - opportunity analysis

The distribution discounting uncertainty analysis reduces species-specific value at locations according to the level of uncertainty there. The approach applies information-gap theory. The level of discounting is specified by the parameter alpha, which in Zonation v.2.0 was constrained to be positive. As alpha is subtracted from species-specific occurrence level, positive values of alpha mean that value is always reduced with growing uncertainty.

In Zonation v.3.0, alpha is no longer constrained, but can be assigned negative values as well. Negative values of alpha can be utilized for an opportunity analysis - search for areas from where a better-than-expected outcome could be obtained. The opportunity variant of uncertainty analysis prefers an area with higher uncertainty to an area with low uncertainty, if the nominal predicted level for the feature (probability of species occurrence) is the same. In other words, increased uncertainty would in this case inflate distribution instead of discounting it. The technical implementation of an opportunity analysis is identical to that of distribution discounting, the only difference being that negative values are assigned to alpha.

Opportunity analysis might be appropriate when one wishes to gain the most from very limited resources. For example, if an area has low value because it is poorly surveyed, then opportunity analysis might reveal that the area could have more value than what is evident based on what is strictly known only. See Ben-Haim (2006) for examples of opportunity analysis.

Literature

Ben-Haim, Y. 2006. *Info-gap decision theory: Decisions under severe uncertainty*, 2nd edition. Elsevier Academic Press, London.

2.5.3 Uncertainty in the effects of landscape fragmentation

This section is mainly based on Moilanen and Wintle (2006).

This analysis is for investigating how robust a given reserve network is to potential negative effects of fragmentation. This is a minor analysis that would not be relevant for most applications of Zonation - essentially, it is intended for the investigation of the robustness of a network that has been designed without any spatial aggregation method. The description below is a brief description of the analysis, please see the original publication for details.

We start from the assumption that cells ending near the edge of a reserve network will end up with decreased conservation value due to negative edge effects, disturbance and disrupted spatial (meta)population dynamics. Just like distribution discounting, fragmentation uncertainty analysis is based on the info-gap theory. Since the main interest of this analysis is uncertainty in conservation value due to habitat loss and fragmentation, error rates (see section 2.5.1) are explicitly related to the amount of habitat loss occurring near the focal cell, assuming the habitat close to the reserve network border is degraded.

The error model chosen for this study specifies that the "true" value of a cell is inside an uncertainty interval;

$$f(\alpha L_c)p^*_{sc} \le p_{sc} \le \frac{p^*_{sc}}{f(\alpha L_c)} \qquad p_{sc} \in [0,1] \qquad (1)$$

where the relative error measure w_{sc} has been replaced by the proportion of cells lost from the neighborhood of focal cell c, L_c. (This is the proportion of cells that were available but were not chosen into the reserve network.) The uncertainty model above is a proportional error model in the info-gap terminology (Ben-Haim, 2006). In the equation, $f(\alpha L_c)$ could be any decreasing function of L_c with $f(0)=1$, and $f(x) \in (0,1)$ for all $x \geq 0$. The condition $f(0)=1$ specifies that if nothing has been lost from the neighborhood of a cell, then the nominal predictions should be used as there is no uncertainty caused by nearby habitat loss. In the Zonation program a simple relationship $f(\alpha L_c)=\exp(-\alpha L_c)$ is used, but any other decreasing function would give similar results. With this relationship, if $\alpha=0$ (no uncertainty) or if there is no habitat loss ($L_c=0$), then from (1) it follows that $p_{sc}=p^*_{sc}$. Increasing the uncertainty parameter α or the fraction of habitat lost will result in a potential loss of conservation value in the target cell.

Note that it is not claimed that any correct value of α is known. Rather, the uncertainty analysis will proceed to analyse how robust different solutions (reserve network candidates) are to increasing uncertainty. The second component needed for the uncertainty analysis is a measure of the performance of a candidate reserve network as uncertainty is increased. The performance of species s in candidate reserve network $\mathbf{X}$, $V_s(\mathbf{X},\mathbf{p})$, was defined as the proportion of the original full distribution of the species remaining in the given reserve structure $\mathbf{X}$;

$$V_s(\mathbf{X},\mathbf{p}) = \frac{\sum_c x_c p_{sc}}{\sum_c p^*_{sc}} \qquad (2)$$

where p^*_{sc} is the original value for species s in the cell c, and p_{sc} is the final value after the loss of neighborhood habitat has been taken into account. x_c receives values depending whether the cell in question is included to the candidate reserve network ($x_c=1$) or not ($x_c=0$). Just like in distribution discounting, also here the evaluation of robustness means that we are interested in the most adverse choices. Therefore, we look at the lower bounds of probabilities. This means that in Eq. (2) p_{sc} is equal to $f(\alpha L_c)p^*_{sc}$.

Different reserve structures will be differentially resistant to negative effects of fragmentation - very fragmented reserve networks will have the characteristic that if species occurring in the network are sensitive to fragmentation, then the true value of the network can be much less than what is estimated based on an analysis that does not include explicit spatial effects. In contrast, a large and well-connected reserve will be relatively insensitive to negative effects of fragmentation, because only a small fraction of the reserve area will be close to the reserve edge. Essentially, such a reserve would have a large core. When running the fragmentation uncertainty analysis, the program calculates how much biological value might be lost from a candidate reserve network as the uncertainty and/or the fragmentation of surrounding habitats increases.

Literature

For more information about uncertainty of fragmentation effects, see

Moilanen, A. and Wintle, B.A. 2006. Uncertainty analysis favors selection of spatially aggregated reserve structures. *Biological Conservation,* 129: 427-434.

For general understanding of information-gap theory, see

Ben-Haim, Y. 2006. *Info-gap decision theory: Decisions under severe uncertainty*, 2nd edition. Elsevier Academic Press.

2.6 Ecological interactions

This section utilizes methods first described in Rayfield et al. (2009), but the method has been applied since in various ways. See e.g. Lehtomäki et al. (2009) and Carroll et al. (2010).

In most cases, conservation planning is done purely based on occurrence data, either of species or other biological features such as vegetation types etc. More specific information is sometimes needed, for example, if we want to include interactions between species and e.g. their food resources, predators or competing species, into the conservation planning process. With Zonation, this type of specific planning is possible as the software includes a facility for modeling a variety of species interactions. In this section we describe the method and philosophy behind this facility and give some possibilities of cases where this type of analysis would be useful.

(1) Interaction type 1 - resource-consumer interactions

Interaction type 1 refers to interactions that are positive for one party but negative for the other. One can use type 1 interactions to account for plant-herbivore, predator-prey, host-parasitoid or present to future interactions.

The general idea is that one wishes to (i) protect the resource independently, (ii) protect a part of the resource distribution which is available to the consumer, (iii) emphasize protection of the consumer at areas that are within foraging distances from the resource. Item (i) can be achieved simply by entering the resource as an independent layer into the Zonation analysis. Items (ii) and (iii) are linked, and they can be implemented via application of the distribution smoothing technique.

Mark by r_{ij} and c_{ik} the local abundances of resource *j* and consumer *k* in grid cell *i*, respectively. Let β_k be the parameter modeling the spatial scale of foraging for consumer *k* . β_k is the parameter of a negative exponential function. We specify that the resource use intensity of resource *j* at cell *i* by consumer *k* is R_{ij},

$$R_{ij} = r_{ij} \max\left\{1.0, \frac{S_{ijk}}{\gamma_j \max_i S_{ijk}}\right\} = r_{ij} \max\left\{1.0, \frac{\sum_{n=1}^{N} \exp(-\beta_k d_{in}) c_{nk}}{\gamma_j \max_i \sum_{n=1}^{N} \exp(-\beta_k d_{in}) c_{nk}}\right\}, \quad (1)$$

which is the local resource density multiplied by the connectivity of the cell to the consumer population, S_{ijk}, using parameter β_k to model the foraging distances of the consumer (d_{in} is distance between cells *i* and *n*). Thus locations with high R_{ij} have both an abundance of resource, and that resource is within the foraging distance to a relatively high number of consumers. Eq. (1) is the connectivity of the resource to the distribution of the consumer. $\gamma_j \in [0, 1]$ is a species-specific parameter describing how fast resource use is saturated. If $\gamma_j=1$, R_{ij} scales linearly with connectivity between distributions, S_{ijk}. If, for example, $\gamma_j=0.1$, R_{ij} achieves its maximum value when connectivity is 10% of the maximum it gets anywhere in the landscape, and so on.

Concerning parameterization, in the two-dimensional case half of the foraging would be performed with a distance of $2/\beta_k$ from the focal cell. Knowledge of foraging distances thus allows setting a reasonable estimate for β_k. Note that the size and unit of the grid cell needs to be accounted for when calculating β_k, see section 3.3.3.8. Essentially β_k, is calculated identically than the scale of landscape use for the distribution smoothing technique.

Eq. (1) is a simple variant of a connectivity computation between distributions. A general variant is

$$R_{ij} = r_{ij} f\left[\sum_{n=1}^{N} \exp(-\beta_k d_{in}) c_{nk}\right], \tag{2}$$

where function *f*() describes an arbitrary (more complicated) function of connectivity, which influences resource use. Notably, in some cases it might be appropriate to use log of connectivity instead of connectivity direct. Such complications are likely to be implemented in future versions of Zonation.

(2) Interaction type 2 - negative interactions

Interaction type 2 refers to interactions that have negative consequences for both parties. The general idea is that one wishes to de-emphasize those parts of the distribution of feature (species) A, that are close (well connected to) the distribution of B. Feature B could be, for example, a competitor, a potential source of an invading species, or a source of pollution that may cause future degradation of habitat quality and consequent reductions in the population sizes of species (biodiversity features) of conservation interest.

Using the notation above, we now specify that the discounted value of feature *j* at cell *i* is R_{ij},

$$R_{ij} = r_{ij}\left\{1.0 - \max\left[1.0, \frac{S_{iju}}{\gamma_j \max_i S_{iju}}\right]\right\} = r_{ij}\left\{1.0 - \max\left[1.0, \frac{\sum_{n=1}^{N} \exp(-\beta_k d_{in}) u_{nk}}{\gamma_j \max_i \sum_{n=1}^{N} \exp(-\beta_k d_{in}) u_{nk}}\right]\right\},$$

(2)

which is the local density of the focal feature r_{ij} discounted by connectivity to the undesirable feature to be avoided *u*, using parameter β_k to model the distances to which the undesirable influence spreads. Effectively, R_{ij} is the distribution of species *j*, which is not connected to the distribution of *u*. In equation (2), the nominator inside the brackets is the connectivity of the focal cell to the distribution *u*. The denominator is the maximum connectivity any cell has to the distribution *u*. Thus, assuming γ_j=1, the fraction term scales from zero (for unconnected locations) to one (for a maximally connected location). If γ_j<1, then the (here negative) effects of connectivity saturate with a lower level of connectivity, as in Eq (1).

Possible analyses with species interactions

Analysis variants that one might imagine doing using the interactions facility include

1. Modeling of predator-prey, resource-consumer and host-parasitoid interactions. In these analyses, the objective would be to ensure protection of those parts of the resource distribution that are “close enough” to be utilized by the consumer. This corresponds to interaction variant 1, in which the connectivity of the resource to consumer is included in analysis. See Rayfield et al. 2009.

2. Application of Zonation to alleviating conservation consequences of climate change. In this analysis the connectivity of the predicted future distribution of the species to the present distribution is used. At least three maps would be entered for each species, present distribution, future predicted distribution and connectivity from future to present. See Carroll et al. 2010.

3. Avoidance of invading species or sources of pollution. Interaction variant 2 can be used to discount a distribution in the proximity to a known or potential source of an invading species. In effect, occurrence levels of the target species will be reduced at locations that are close to (well connected to) the distribution of the invasive species (or source of pollution).

4. Modeling of food chains or food webs. The interactions can be calculated between multiple species in the analysis, multiple resources can be connected to one consumer etc. By appropriately chaining connectivity effects between distributions, it should be possible to model more complicated relationships than just simple pair-wise interactions.

5. Emphasizing proximity to existing reserve areas when planning expansion of reserve network (Lehtomäki et al. 2009).

When Zonation is started, the species and uncertainty layers are read in one by one. For each, distribution discounting is done first, and then distribution smoothing (if used). Interactions are implemented after all species layers have been read and discounted/smoothed. The interactions are read in one by one and performed between the layers specified in the interaction definitions file. Each interaction transforms one of the species layers that were read in. This means that if, for example, a habitat quality layer, a connectivity layer and interaction layer are to be calculated and used based on the distribution map of one species, then the distribution map needs to be entered three separate times into Zonation: once as a plain unsmoothed habitat quality layer, a second time smoothed to implement the connectivity calculation and a third time to be used in the interaction as a focal layer. Note also that a species map may be transformed by more than one interaction. If the same layer is the focal layer of an interaction multiple times, then the layer models simultaneous connectivity to multiple (different) sources.

After interactions have been implemented, Zonation proceeds as before. Connectivity methods BQP and BLP operate as before, also on interaction layers. Instructions for including species interactions into Zonation can be found in section 5.1.9.

Literature

The method to include considerations of ecological interactions was first described by

Rayfield, B., Moilanen, A. and Fortin M.-J. 2009. Incorporating consumer-resource spatial interactions in reserve design. *Ecological Modelling*, 220: 725-733.

See also

Carroll, C., Moilanen, A., and J. Dunk. 2010. Designing multi-species reserve networks for resilience to climate change: priority areas for spotted owl and localized endemics in the pacific North-West USA. *Global Change Biology*, 16: 891-904.

Lehtomäki, J., Tomppo, E., Kuokkanen, P. Hanski, I., and A. Moilanen. 2009. Planning of forest conservation areas using high-resolution GIS data and software for spatial conservation prioritization. Forest Ecology and Management, 258: 2439-2449.

2.7 Replacement cost analysis

Very seldom can reserve selection start with a clean sheet so that the planning region would have no existing reserves nor any restrictions for conservation (such as areas ear-marked for residential building). More commonly, due to logistic or social constraints, certain sites need to be included to, or excluded from the final solution. In most cases this leads to a suboptimal network either in terms of conservation value or in terms of the economic cost of achieving a given conservation goal (compared to the one that could have been obtained with a clean sheet start).

It is useful to be able to asses the degree of suboptimality of solutions compared to the optimal ones. A method called *replacement cost analysis* can be used to evaluate the effects of forced site inclusion/exclusion. *Replacement cost* refers to **loss in the value of the solution** given that the optimal cost-efficient solution cannot be had and that alternative solutions, with particular sites forcibly included or excluded, must be accepted. It tells us the cost (biological or economic) of including or excluding a site from the reserve network. If budget is constant, the *exclusion cost* of a site is the loss in the network's conservation value that follows when the site belongs to the optimal solution but has to be excluded from the reserve network. The *inclusion cost* of a site is the loss in conservation value that must be accepted if a suboptimal site is forced into the reserve network. On the other hand, when the conservation budget is not fixed, replacement cost can also be defined in terms of the extra funding required to maintain conservation value that is equal to the value of the optimal solution. Thus, one can define four variants of replacement cost:

1. *Biological exclusion cost.* Decrease in conservation value following forcible exclusion of a given (group of) site(s).

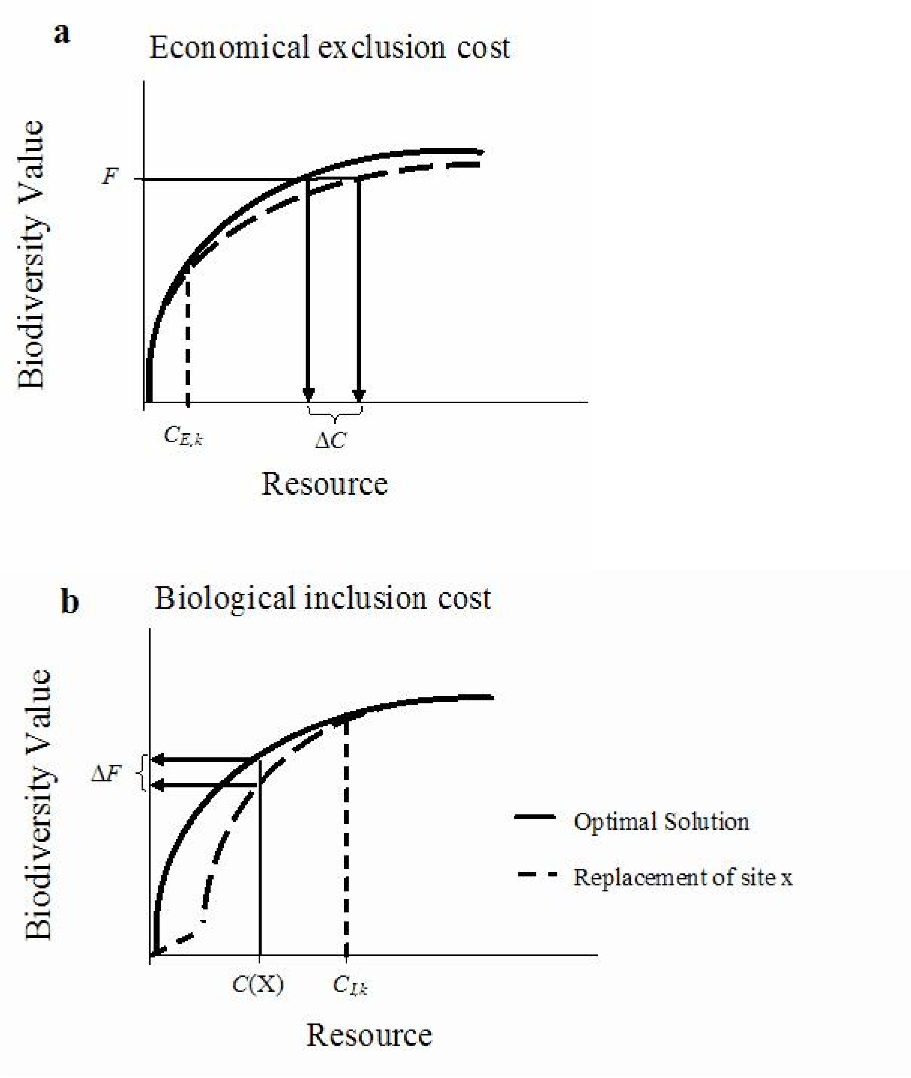
2. *Biological inclusion cost.* Decrease in conservation value following forcible (suboptimal) inclusion of a given (group of) site(s).

3. *Economic exclusion cost.* Increase in solution cost required to keep the same total conservation value following the forcible exclusion of a given (group of) site(s).

4. *Economic inclusion cost.* Increase in solution cost required to maintain conservation value following the forcible inclusion of a given (group of) site(s).

For practical purposes, the replacement cost is calculated in the following manner. First, find optimal reserve selection **X*** which has highest possible value *F*(**X***) obtainable with the available resource (C_{max}). Then rerun the analysis with some areas forced in or out. Replacement cost is the difference between the value of the optimal solution and the value of the new solution. We emphasize that **this does not mean keeping the optimal set of sites plus/minus a particular site**, **but finding a completely new solution given that the particular site(s) are forcibly included/excluded.**

A replacement cost of zero tells us that there exists an alternative solution with the same value as the current (best) solution has, i.e. same cost, and same conservation value (although obtained via a different selection of areas compared to the original optimal selection). A replacement cost larger than zero means that any alternative solution including/excluding the focal site(s) will have either a lower conservation value or a higher economic cost than the optimal one.



A conceptual illustration of the *replacement cost* of a hypothetical site (a) in terms of increase in resources required to maintain value (Δ*C*) and (b) in terms of loss of biodiversity value (Δ*F*) (from Cabeza and Moilanen 2006). (a) *Exclusion cost*: the dashed line indicates the value of the best solution when site *x* is forcibly excluded. Up to a certain resource level, $C_{E,k}$, site *x* does not belong to the optimal solution and thus exclusion cost is zero. Even with $C > C_{E,k}$, exclusion cost can be zero if the site is fully exchangeable with another site or a combination of other sites. (b) *Inclusion cost*: the dashed line indicates the value of the best solution when a site is forced to be included in the solution. Inclusion cost is likely to be highest with low resource, when the forced inclusion of the unwanted site prevents the acquisition of other biologically much more valuable locations. At $C > C_{I,k}$, the focal site becomes included in the optimal set and inclusion cost becomes zero. (Note that the resource here should be understood as the proportion of landscape retained in the Zonation.)

The figure also shows an example of how exclusion and inclusion can be expected to behave qualitatively. With a small resource ($< C_{E,k}$), the exclusion cost (a) of a site is likely to be zero, because the site would not be in the optimal set in any case. At a level $C_{E,k}$ the site becomes part of the optimal solution. With a resource slightly higher than $C_{E,k}$, it is likely that the exclusion of the site can be compensated with small cost, at least if there are many selection units. However, when the available resource is large, sites of less importance are included in the solution and the exclusion of a high-quality focal site has a clearly positive cost.

Inclusion cost (b) behaves differently. When a site that would not belong to the optimal solution is included in the network it generates an increase of cost even when the resource available is small because the resource is spent on suboptimal areas. With increasing resource availability, the inclusion cost gradually decreases. At a level $C_{I,k}$, the site would already become part of the optimal solution and inclusion cost thus becomes zero.

Instructions for including and excluding areas to a Zonation solution for analyzing replacement cost can be found in section 5.2.3.

Literature

Replacement cost analysis is described by

Cabeza, M. and Moilanen, A. 2006. Replacement cost: a useful measure of site value for conservation planning. *Biological Conservation,* 132: 336-342.

See also

Leathwick, J. R., Moilanen, A., Francis, M., Elith, J., Taylor, P. Julian, K. and T. Hastie. 2008. Novel methods for the design and evaluation of marine protected areas in offshore waters. *Conservation Letters*, 1: 91-102.

Moilanen, A., Arponen, A., Stockland, J.N. & M. Cabeza. 2009. Assessing replacement cost of conservation areas: How does habitat loss influence priorities? *Biological Conservation,* 142: 575-585.

2.8 Community level analysis

This section follows the manuscript in which community-level analysis in Zonation was first described:

Leathwick, J.R., A. Moilanen, S. Ferrier and K. Julian. 2010. Complementarity-based conservation prioritization using a community classification, and its application to riverine ecosystems. *Biological Conservation*, 143, 984-991.

As legislation and politically set conservation targets often consider habitat or community types, it is often reasonable to consider these when selecting conservation priorities. Considering representation of community types instead of individual species can also lend itself in situations in which information about individual species distributions is too sparse for fitting reliable habitat suitability models.

Zonation can utilize (i) grid layers that define distributions of community types and (ii) a similarity matrix that describes pair-wise similarities in species composition (or in some other aspect). (iii) The third component used in community level analysis is information about the species richness of each community class. Species richness is subsumed into the weight given to the community class (or habitat type), see Leathwick et al. (2010) for details. Typically, this analysis would utilize a (statistically-based) community type classification, in which the effective number of community classes is reduced to some tens or hundreds.

The pair-wise similarities can be obtained in a number of ways. An ecologically reasonable method for quantifying similarities in species composition is generalized dissimilarity modeling (GDM; Ferrier 2002). A GDM models turnover in species composition as a function of environmental variables. In a way, it calibrates environmental gradients to match the turnover rate in species composition at different segments of the gradient. A GDM can then be used to predict compositional dissimilarities between two sites, for which the relevant environmental variables are known (see Ferrier et al. 2007).

Zonation uses the similarity matrix to expand the primary habitat type classifications across the landscape to effective occurrences, which also account for similarity between habitat types. The fraction f of a community type j protected under management actions **a** (in this context **a** would most often be reserving areas for conservation) is defined here as

$$f_j(\mathbf{a}) = \frac{\text{protected}_j}{\text{total}_j} = \frac{\sum_{k=1\ldots n}\sum_{i=1\ldots m} L_{ik} S_{ij} C_{ik}(a_k)}{\sum_{k=1\ldots n}\sum_{i=1\ldots m} L_{ik} S_{ij}}$$

where L_{ik} describes the extent of classification group i in cell k, S_{ij} describes the modelled biological similarity between groups i and j, and C_{ik} describes the landscape condition (see section 2.10.) for group i in cell k. The effective occurrence level for group i is, then, the sum of the occurrence level x similarity of all other groups present in cell k.

If, for example, community type A has 40% of its species in common with community type B, cells where type B is present (with occurrence level of 1) is considered to represent community type A with an occurrence level of 0.4. Also, the original community type data do not need to be binary presence/absence, but they can reflect e.g. the proportional cover of the community type in the grid cell.

1	0	1
0	0	1
0	1	0

Community type A

0	1	0
1	1	0
0	0	0

Community type B

1	0.4	1
0.4	0.4	1
0	1	0

Community type A after similarity expansion

A figure illustrating the similarity expansion. Here, 40% of species characteristic to community type A are also present in community type B.

As Zonation prioritization takes place only after the similarity expansion, it is important to understand that representation of community types in the Zonation solution does not exactly correspond to the original community classes, but the expanded classes. A representation level of 0.4 of community type A in a given top fraction does not mean that 40% of the original community type A is covered in the solution, but rather that a fraction 0.4 of effective occurrences are covered. It is, however, easily possible to return to original unexpanded classifications and check their coverage in any selected area of a solution.

Note that community-level analysis would typically be combined with at least condition information and possibly retention. See the matrix connectivity computations for connectivity that is applicable to a set of partially overlapping habitat types (section 2.4.5.; Lehtomäki et al. 2009).

Instructions for using community level biodiversity features are in section 5.3.1.

Literature

Community level analysis has been applied by

Leathwick, J.R., A. Moilanen, S. Ferrier and K. Julian. 2010. Complementarity-based conservation prioritization using a community classification, and its application to riverine ecosystems. *Biological Conservation*, 143, 984-991.

NB. Leathwick et al. did not use similarity expansion of Zonation v.3.0., but pre-processed data using GIS and R. Following their protocol, it is possible to use the technique also with Zonation v.2.0.

Lehtomäki, J., Tomppo, E., Kuokkanen, P. Hanski, I., and A. Moilanen. 2009. Planning of forest conservation areas using high-resolution GIS data and software for spatial conservation prioritization. Forest Ecology and Management, 258: 2439-2449.

For more information about generalized dissimilarity modeling, see

Ferrier, S. 2002. Mapping spatial pattern in regional conservation planning: where to go from here? *Systematic Biology* 51: 331-363.

Ferrier, S., Manion, G., Elith, J. and Richardson, K. 2007. Using generalized dissimilarity

modeling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions* 13: 252-264.

2.9 Alternative land uses - multi-criterion analysis

The primary function of Zonation is to identify priority areas for conservation. Its purpose is thus not to produce comprehensive land use plans for zoning, for example. Conservation is, of course, not independent of other land uses, and therefore it is useful to take multiple land use criteria into account in the conservation planning process. The ability to balance between biodiversity and competing land uses is relevant when considerations of cost must influence decisions or when compromises between conflicting interests need to be sought. Previous versions of Zonation could be adjusted to mask out areas that are out of question for conservation, such as heavily constructed urban areas or power plants, industrial harbors etc. It was also possible to incorporate variation in land cost into the analysis. Zonation v.3.0. goes a step further and allows balancing of multiple, possible antagonistic, land use criteria in prioritization. The idea is to find a spatial separation between positive biodiversity features and negative alternative land uses. Biodiversity should be retained into the top fraction of the priority ranking whereas areas suitable for alternative (negatively weighted) land uses should receive low priority ranks.

Technically, any analysis with more than one feature optimized is multi-criterion analysis, even when there are only two biodiversity features. Here, however, the term multi-criterion analysis is reserved for the case that there are biodiversity considerations and opposing land uses; opportunity costs etc.

Including competing land uses into Zonation is remarkably simple. Competing land uses are fed into Zonation in a similar way as distribution layers of species or other biodiversity features. The difference is that the competing land use features are assigned negative weights (sections 3.3.2.2. and 5.1.1.) instead of positive ones. This will enable Zonation to give high conservation priorities to sites that have high conservation value but are not very favorable for alternative land uses. Features with negative weights are preferably removed early on in the prioritization, while biodiversity is retained.

The structure of conservation value in Zonation now becomes
(biodiversity - competing land uses) / cost, where cost should be taken as the direct cost to the conservation agency.

There is not rule as to how big negative weights should be used relative to the positive weights for beneficial biodiversity features. If it is critical that almost all of the negative feature remains outside conservation areas, then the negative weight should be relatively large. A small negative weight is appropriate for mildly detrimental features. Experimenting with the relative weights will allow finding a balance between positive and negative features that answers planning needs.

The difference between forcing cells to be removed first with a removal mask (section 3.3.3.9.) and multi-criterion analysis is that removal mask will discard the areas completely, whereas negative weighting of competing land use features allows the sites to still be taken into account in the planning process. Multi-criterion approach is more suitable, when suitability of sites for competing land uses is varied, and identifying solutions to account for several purposes are needed. Excluding areas completely already at the starting point with exclusion mask is more appropriate if it is already known that some sites will be used in ways that do not support biodiversity.

Please note: While it is useful to take multiple criteria into consideration in the conservation planning process, the primary focus in conservation prioritization should be maximizing biodiversity value. If too much attention is given to cost and other secondary matters, then they may end up being the main drivers of conservation decisions. This would reset systematic conservation planning back to its starting point, in which protected area networks were not representative precisely because unexploitable or otherwise least costly areas were designated to conservation (for discussion see Arponen et al. 2010).

Instructions for including considerations of multiple land uses are in section 5.3.4.

Literature

Multi-criterion analysis using Zonation has first been applied in
Moilanen, A., Anderson, B.J., Eigenbrod, F., Heinemeyer, A., Roy, D.B., Gillings, S., Armsworth, P.R., Gaston, K.J. and Thomas, C.D. Balancing alternative land uses with the Zonation conservation prioritization approach. *Ecological Applications*, 21: 1419-1426.

For discussion about cost considerations in conservation planning, see
Arponen, A., Cabeza, M., Eklund, J., Kujala, H. and Lehtomäki, J. 2010. Costs of integrating economics and conservation planning. *Conservation Biology* 24: 1198–1204.

2.10 Landscape condition and retention

Landscape condition and retention are important features in their own right, and integral features of community level analysis (section 2.8.).

Each species or other biodiversity feature can be linked to a condition and retention group. This means that certain types of habitats and groups of species have specific responses to prior and expected habitat loss. These responses need not be, and generally would not be, specific to feature. Rather, the responses would be specific to groups of features, such as community types or guilds of species.

(i) Condition

In Zonation, condition represents information about local habitat deterioration and its influence on species (features) or groups of features. Condition is specific to species (group) and grid cell - typically, there are several condition map layers that can be linked to different groups of features.

The value of landscape condition can vary between 0 (all local conservation value has been lost) and 1 (all habitat remains locally in a pristine state; well, compared to some historical baseline). The influence of condition in analysis is the following: condition normalizes present landscape condition to be measured against a historical baseline. The point is that remaining occurrences for features that have lost a lot are relatively more valuable than occurrences of features that have lost little. For example, habitats (and their species) occurring on fertile low-land soils may have suffered from significant losses. Low-productivity habitats occurring at high elevations and rough terrains would have suffered little loss. Use of condition allows one to model this past loss in a feature-specific manner.

(ii) Retention

This section follows closely the following manuscript in which the use of combined condition and retention in spatial prioritization has first been described:

Moilanen, A., Leathwick, J.R. and Quinn, J.M. 2011. Spatial prioritization of conservation management. *Conservation Letters*, 4: 383-393.

The conventional systematic conservation planning process aims at maximizing representation of biodiversity features in a network of protected areas. It carries the implicit assumption that areas outside the network contribute nothing to the overall biodiversity value of the region. This may be a realistic assumption if pressure for land conversion is high. However, often conditions are such that habitat is retained in the landscape even when some areas are not protected. This may be the situation with areas that are remote and difficult to access, or with areas that are not economically exploitable. In such situations, also the non-protected areas will support biodiversity and this could and should be accounted for in the conservation planning process. Zonation v.3.0 can account for the occurrences of features that are retained across the whole landscape. There are two ways in which landscape can be retained: a) retention implies that not all conservation value is lost from a location even without conservation action, or b) management intervention increases habitat quality or population sizes at retention sites.

Most often one would like to account for both representation in the protected area network and retention across the entire landscape. Use of only representation can hamper effective conservation, as some biodiversity features are likely to be well retained also in non-protected areas. Then again, relying only on retention is risky, as high uncertainty is inherent in assumptions about threat and loss of habitat. To consider both representation and retention, Zonation needs two sets of biodiversity feature layers (in practice, this means that you need to list twice the set of biodiversity features you want to use for the analysis). For one set, representation, it is assumed that all conservation value is lost if the site is not protected. The other set of layers contain the difference between what is expected to remain if the site is protected and what will remain even in the absence of conservation action. The relative weighting of these sets of layers will determine how much emphasis is placed on representation and how much on retention. A high weight on retention layers means that emphasis is placed on locations where conservation action can make a difference. A high weight on representation means emphasis on locations where rarity and richness of features is highest based on known data.

Operation of Zonation, when using condition and retention

When Zonation reads in raster grid files, the following steps happen:

(1) You read in absolute occurrence levels for features *j* in cells *i*. Call these O_{ij}.

(2) O_{ij} is normalized to unity (1.0) by division with sum of O_{ij} across cells *i*. Call this normalized occurrence N_{ij}; ? N_{ij}=1.0. At this stage also connectivity transforms such as distribution smoothing and matrix connectivity are done for selected layers, after which the distributions are renormalized.

(3) Apply condition to each cell for each relevant feature. => N_{ij}(condition)= N_{ij} * C_{ij}, where C_{ij} is the fractional quality read in from the respective condition group.

This is a condition transformed distribution for the feature, and it no longer sums to unity. So far everything is straightforward, *but in the following steps it is relevant to pay attention to the meaning of retention matrixes at input*.

(4) Retention is applied on top of condition. Retention is applied to calculate the difference made with and without conservation - retention ranking is based on the difference between what happens with and without conservation. At input, retention for a cell, R_{ij}, is first transformed to fractional DIFFERENCE MADE with respect to N_{ij}(condition). We state that the fractional difference D_{ij}=abs(1- R_{ij}), where the absolute value returns a positive difference independent of whether we use mode 1 (R_{ij}<1) or mode 2 (R_{ij}>1). A retention value of 1.0 implies no change in condition even in the absence of conservation action.

(5) For stop loss (retention mode 1), the assumption is that condition goes *down* by the difference in the absence of action.

(6) For management gain (retention mode 2), the assumption is that condition goes *up* by the difference in the absence of action.

(7) The difference between stop loss and management gain manifests in that the weight for the retention transformed layer is calculated in a different manner, as described next.

(8) Computation of weights for retention-transformed features

The set of features used for representation can already have variable weights, e.g. following a species weighting scheme or because different communities inherently differ in species richness. These weights are denoted by w_j for community type *j*. Marking by w^r_j the weights given for retention-transformed layers, one could assume that $w^r_j = w_j$, but this would not be sensible, as different community classes are likely to be losing different fractions of their remaining distribution. A high weight should go to a feature that would in the absence of conservation action lose a large part of its remaining distribution. When the expected relative loss for the feature would be small, the weight for the respective retention-transformed layer should be small. Values for w^r_j come out different for the two analysis modes:

(8a) Feature weight calculation for retention Mode 1, stop loss

The weight calculation is linked to feature-specific occurrence levels, condition and retention for each grid cell in the landscape. To correct for variable expected loss in the absence of conservation, we set

$$w_j^r = \beta \frac{difference_j}{remains_j} w_j ,$$

where parameter β tunes the balance between representation and retention in the analysis; the larger the value of β, the more emphasis is given to retention. Both *difference* $_j$ and $remains_j$ are measured for feature *j* across the full landscape. $Remains_j$ is fraction of occurrences remaining for the feature after application of condition:

$$remains_j = \sum_i c_{ij} N_{ij} ,$$

in which N_{ij} is normalized pristine state (or other past baseline) occurrence level of feature *j* at grid cell *i*, and c_{ij} is condition of feature *j* in grid cell *i*. $Difference_j$ is what fraction of $remains_j$ would be lost in the absence of conservation:

$$difference_j = \sum_i abs(1 - R_{ij}) c_{ij} N_{ij} ,$$

where R_{ij} is the respective retention value.

(8b) Feature weight calculation for retention Mode 2, management gain

Using the notation above, we now set

$$w_j^r = \beta \frac{difference_j}{difference_j + remains_j} w_j ,$$

where the point is that with application of management intervention there is a gain, and the relative loss is from better than present condition to present condition.

Instructions for using landscape condition and retention in your analysis are in sections 3.3.3.14., 3.3.3.15. and 5.3.5.

Literature

Landscape condition in Zonation has been applied by
Leathwick, J.R., Moilanen, A., Ferrier, S. and Julian, K. 2010. Community-based conservation prioritization using a community classification, and its application to riverine ecosystems. Biological Conservation, 143: 984-991.

For discussion on retention, see

Moilanen, A. and M. Cabeza. 2007. Accounting for habitat loss rates in sequential reserve selection: simple methods for large problems. *Biological Conservation*, 136: 470-482.

Pressey, R.L., Watts, M.E. and T.W. Barrett. 2004. Is maximizing protection the same as minimizing loss? Efficiency and retention as alternative measures of the effectiveness of proposed reserves. *Ecology Letters* 7: 1035-1046.

2.11 Landscape dynamics

Another thing that can be done with increased memory is faking of landscape dynamics. For this analysis, habitat suitability maps for a set of species or other biodiversity features are entered for the present and for several time steps in the future. This fakes dynamic landscapes and requires solutions that are balanced at all time steps. This kind of brute force approach really only is possible with the increased memory capacity of Zonation v.3.0. Connectivity between time steps could in this analysis be modeled by some connectivity method which accounts for multiple distributions simultaneously. In Zonation, this would be interaction connectivity (section 2.6.) or matrix connectivity (section 2.4.5.).

2.12 Administrative units

This section follows closely a manuscript in which the method was first described:

Moilanen, A. and A. Arponen. Administrative regions in conservation: Balancing local priorities with regional
to global preferences in spatial planning. *Biological Conservation*, 144: 1719-1725.

In reality, conservation decisions are usually taken at national or regional levels, or even at the scale of individual land parcels. Distributions of species and other biodiversity features are spread over multiple such administrative units. Different administrative units may, for whatever reasons, have different priorities for conservation of biodiversity and its components (Arponen et al. 2005). As population dynamics and connectivity effects do not respect such administrative borders but extend across them, it is reasonable that the global conservation status of a biodiversity feature should influence its conservation locally. How should the global and local conservation needs of a biodiversity feature then be balanced to account for connectivity across borders?

It is possible to account for variable local and global priorities in the Zonation analysis via locally variable weights that are assigned to species or other biodiversity features. We describe below two alternative methods for doing so. The first relies on what we call weak local representation. The emphasis with this approach is in global representation with locally varying weights; it allows a degree of flexibility between what features are represented in which administrative regions. The second option, strong local representation, requires all features to be represented separately in each administrative region, when at all possible. This is irrespective of how the local abundance of the feature compares to global abundance.

Weak local administrative priorities (ADMU mode 1)

In this analysis variant it is assumed that priorities (feature weights) can vary between administrative regions, but that analysis is primarily global, and that a degree of substitution of representation between regions is allowed. For example, if species *j* is rare in area *A* but common elsewhere, representation of *j* within *A* would not be enforced. Representation for feature *j* is preferably obtained from a location where both the priority of *j* is high and the feature occurs at high local occurrence levels. Thus, there is no explicit guarantee that all features would become "protected" in all regions.

Technically, the implementation of this analysis variant in Zonation is simple. The equations that define the marginal loss of conservation value following the removal of a focal cell from the conservation solution include feature-specific weights, w_j (Moilanen et al. 2005, 2008; Moilanen 2007). Here, in addition to the global feature weights, w^G_j, the features can be assigned different weights for each administrative region separately, w^L_{jA}. In other words, one uses weight w^L_{jA} for feature *j* if the focal grid cell is part of administrative region *A*. Zonation calculates effective feature- and area-specific weights w^{eff}_{jA} by combining global and local weight given to features.

The global loss of benefit is defined as

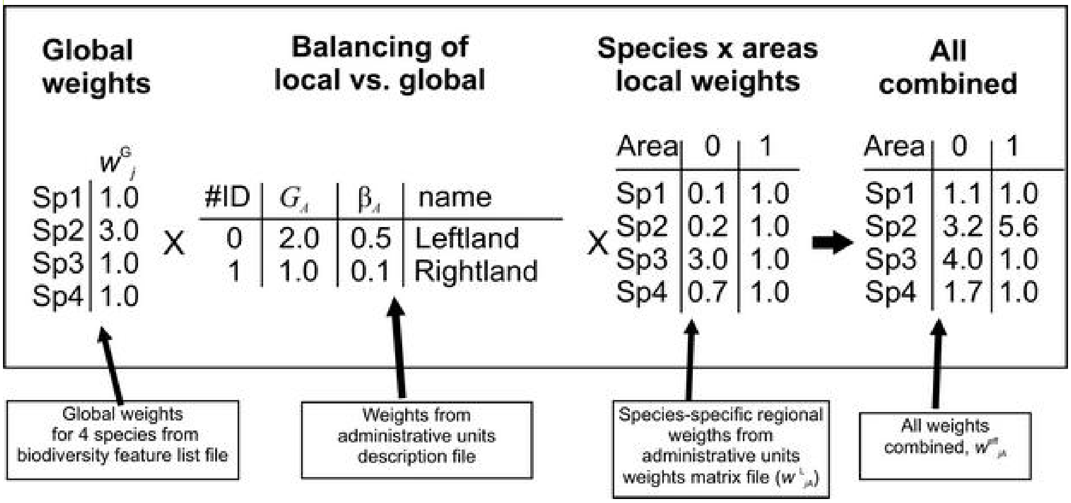
$$\Delta^{weak}(\mathbf{S},i)=\sum_j w^{eff}_{j,A_i}\left\{f_j\left[R_j(\mathbf{S})\right]-f_j\left[R_j(\mathbf{S}-\{i\})\right]\right\}.$$

where the value is converted via the feature-specific benefit function f_j and multiplied by the effective local weight of the feature. Thus, in this analysis, representation is global while effective weights are local.

The effective local weights are constructed as a composite of global and local considerations in the following manner,

$$w^{eff}_{jA}=\left[\beta_A w^G_j+(1-\beta_A)w^L_{jA}\right]G_A,$$

in which w^{eff}_{jA} is the effective local weight, w^G_j is the global weight of feature j, w^L_{jA} is the locally given weight of feature j in area A, and β_A is a parameter tuning how much weight administrative area A places on global considerations. An administrative region can decide to focus only on global priorities (β_A=1) or completely ignore global priorities (β_A=0), or anything between ($0 < \beta_A < 1$). G_A is a region-specific weight which is used to multiply the species-specific weighting within each region. The figure below describes an example of how the effective local feature weights are constructed.


Global weights
Balancing of local vs. global
Species x areas local weights
All combined
X
X
Global weights for 4 species from biodiversity feature list file
Weights from administrative units description file
Species-specific regional weigths from administrative units weights matrix file (w^L_{jA})
All weights combined, w^{eff}_{jA}

	w^G_j
Sp1	1.0
Sp2	3.0
Sp3	1.0
Sp4	1.0

#ID	G_A	β_A	name
0	2.0	0.5	Leftland
1	1.0	0.1	Rightland

Area	0	1
Sp1	0.1	1.0
Sp2	0.2	1.0
Sp3	3.0	1.0
Sp4	0.7	1.0

Area	0	1
Sp1	1.1	1.0
Sp2	3.2	5.6
Sp3	4.0	1.0
Sp4	1.7	1.0

Strong local administrative priorities (ADMU mode 2)

Heuristically, this analysis variant proceeds from the assumption that all locally occurring biodiversity features must be represented locally, but that global considerations should nevertheless influence local priorities. This differs from weak local priorities in that (instead of only weights) weights, representation and conservation value are all considered both globally and locally. Conservation value is aggregated from a global component together with a local component for each administrative area A. Representation of each feature is considered separately globally, and in each administrative region. Effectively, the analysis thus joins multiple conservation prioritization analyses, one global, and one per each administrative region.

This is how conservation value is aggregated in the case of strong local representation:

$$V^{strong}(\mathbf{S}) = p\sum_{j} V_j^{global}(\mathbf{S}) + (1-p)\sum_{A}\sum_{j} V_{jA}^{local}(\mathbf{S})$$

$$= p\sum_{j} w_j^G f_j^G\left[R_j(\mathbf{S})\right] + (1-p)\sum_{A} G_A \sum_{j} w_{jA}^L f_{jA}\left[R_{jA}(\mathbf{S})\right].$$

The analysis is performed by a global actor who has the possibility to decide how much weight is given to priorities of local administrations and how much to global considerations by adjusting the balancing parameter *p*. When *p* is zero, the analysis is entirely based on priorities within each local administration (right hand side of the equation), and when *p* is one the analysis corresponds to a purely global analysis without any consideration of priorities of the local administrations. The priorities of each local administration are on the right hand side of the equation. G_A is an additional region-specific weight, which determines the region's value with respect to other regions. If a region's G_A is 10, then all value within the region, derived from the combined global and local feature weights and respective benefit functions, is multiplied by ten. Also, conceptually, one could use different benefit functions for a feature globally and in each administrative region, as allowed by different functions f_j^G and f_j^A, but this is currently not implemented in Zonation. Representation also has to be treated both globally, R_j(**S**), and locally R_jA(**S**), where the latter would be the fraction of the local occurrence of *j* covered in area *A* under some conservation solution **S**. Parameter β_A from the ADMU descriptions file is not used in the strong priorities variant.

The instructions for running a Zonation analysis accounting for administrative regions are in section 5.3.8.

2.13 Assumptions & limitations

This section lists some known assumptions and main limitations of the presently available Zonation implementation.

- The Zonation software presently only accepts data in grids and point observation lists, in particular, it does not accept vector-based planning units. This limitation is not practical, not conceptual, but it is unlikely to be removed in the next Zonation version.

- Zonation is presently for doing an implement-in-one-go reserve/conservation plan. In particular, it does not include any explicit mechanism for handling multi-year planning with considerations of stochastic site availability and the possibility of site loss. The retention analysis does however account for the first-order effects of expected habitat loss.

- At present Zonation only has a single option per selection unit, that is, protect or not, restore or not, maintain or not and so on. In a more advanced analysis one could envision multiple alternative outcomes for each location (for example, different levels of protection or restoration). However, one can do analysis where alternative conservation options are explored. This involves developing sensible restoration scenarios and then running Zonation prioritization for each (see Thomson et al. 2009).

- Zonation v.3.0 includes many different feature-specific forms of connectivity, but it does not operate on path-like connectivity. This deficiency is likely to be corrected in forthcoming versions of Zonation.

- The Zonation software does not include a full GIS interface and only a limited set of analyses and graphs can be produced with the software. The analysis output files can however be imported into GIS software for further processing.

Major limitations removed since Zonation v.2.0

- Multi-criterion analysis and balancing of alternative land uses is now possible

- Ability to work with environmental data and community classes

- The 4GB 32 bit Windows memory limit

- Effects of conservation action can now partially be accounted for via retention analysis.

Part III

3 ZIG - The Zonation software

3.1 Important general information about file operation

In the following sections we describe how the Zonation software is used from the command line, what kind of input files it requires, what kind of output it produces and which are the analyses that can be done with the program. For concepts and analyses implemented in the software, see section 2. In Zonation v.3.0 there is a new Graphical User Interface. It is a separate piece of software which calls zig3.exe executable to execute analyses. We describe the use of the new GUI in section 4.

File names

The length of file names in Zonation v.3.0 is not limited and they may contain numbers and any unicode characters. File names containing special characters such as spaces can be read if they are in quotation marks: e.g. "my zonation run.bat". Directory separators can be entered as slashes or backslashes. File names can be enclosed in double quotation marks. In this case, quotation marks in paths can be entered in escaped form. For example, the file path c:\my “great” zonation run.bat can be entered as “c:/my \”great\” zonation run.bat” in quoted form.

Specifying file paths

When you enter file names, please note that **you need to specify the path to your file** as well. In the command prompt, the paths in text files are relative to the current working directory. In the GUI they are relative to the project batch file. If all your files are in the same folder, the file name will suffice:

```
my_file.asc
```

However, most often it is convenient to have subfolders for different analyses. If the subfolders partition from the folder where you keep the batch files, please specify the path from that folder:

```
subfolder/my_file.asc
```

or with the full path:

```
D:/Data/subfolder/my_file.asc
```

This applies to all files that you refer to at any stage of your analysis setup. In the examples, we have left the paths out for the sake of simplicity.

Supported file formats

Zonation uses a modified version of GDAL 1.8.0 (Geospatial Data Abstraction Library) to read and write raster files. In principle, all formats that are compiled by default are supported, but only Arc/Info ASCII Grid (.asc), Erdas Imagine (.img) and GeoTIFF (.tif) are tested. The complete list of supported formats can be found at http://www.gdal.org/formats_list.html.

Any supported file format can be loaded anywhere in Zonation where a raster file is needed. Different formats can be used simultaneously. For example, a part of the species files can be in ASCII format while the rest are in GeoTIFF. If a multiband dataset is used in Zonation, only the first band is read and the rest are ignored.

We recommend the use of compressed raster formats with Zonation. Very commonly the used datasets are masked or mostly zero which makes the compression very effective. It is not uncommon to have an ASCII based dataset squeezed to less than a hundreth of the size of the original when transformed into a deflated GeoTIFF. It is also much faster to read and write smaller files. With dataset sizes in tens or hundreds of gigabytes this can mean big savings.

All text files are expected to be encoded in utf-8.

3.2 Running Zonation

The Zonation numerical core operates through a command line application. You have two options for managing Zonation operations: you can call Zonation directly on command line (section 3.2.1.) or use the Zonation GUI (section 4) to manage and monitor your runs. The GUI for Zonation v3.0 is fundamentally different from the Windows interface of the previous version. It is useful for setting up and managing multiple analyses for a single set of input data. The GUI can also attach to runs that have been started from a batch command and display their progress.

With Zonation you can also either (i) make new analyses or (ii) loading old solutions (section 3.2.3.). The solution loading option can be used either to review old solutions or to investigate how old solutions would perform under new (different) assumptions (section 3.5.3.).

After running Zonation, please check the .run_info.txt -file (section 3.4.1.) for verification about data characteristics and options used, and about any warnings or errors that occurred during the run - the present Zonation version does not stop execution at every suspicious input.

3.2.1 Command line

A practical way to run Zonation is calling it with a batch file. A batch file is a simple Windows/DOS command line file that can be used to give commands to Windows. Batch files can be created for example with notepad. This is done by writing the program call into a new notepad document and saving it with the **batch file -extension** (.bat). The file name extension of a batch file has to be **.bat**, e.g. do_zig.bat. If there is anything else after the .bat -suffix, Windows cannot identify the file as a command file. Batch files are practical because:

- The information of your analysis (which input files and settings have been used) will be saved in the batch file and you can review it later.
- With batch files you can run multiple analyses in one go.

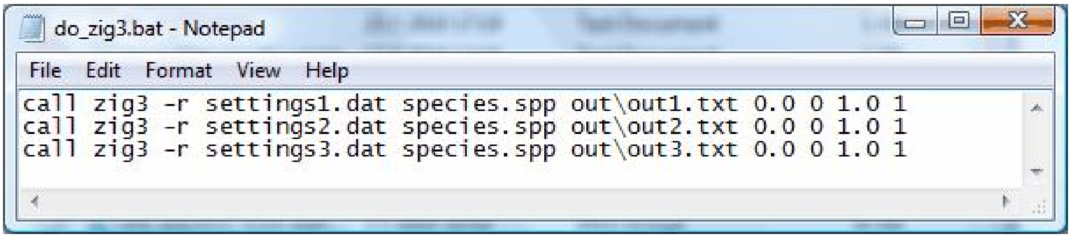
NOTE: Because with Zonation v.3.0 the DLL libraries need to be located within the same directory as the .exe, it is no longer advisable to make a new copy of the program to the directory containing data files when starting a new project. Instead, it is more convenient to establish a permanent directory for the executable and the DLL files and to call the program from that directory, e.g. "call C:\Zonation\zig3.exe".

Call the program by typing "`zig3`" in to your batch file, followed by:

- -r, if a new solution will be calculated, or -lfilename if an existing solution is loaded (the name of the rank.asc -file, e.g., -`loutput.rank.asc`)
- Name of the run settings file (.dat; section 3.3.2.3.)
- Name of the biodiversity feature list file (.spp; section 3.3.2.2.)
- A name for the output files (`subfolder\my_output_name.txt`).
- Value of the uncertainty parameter α for the uncertainty analysis (UCA; section 5.1.8.). If uncertainty is not included in the computation, this should be set to zero.
- Value to determine whether distribution smoothing (section 5.1.3.) is used in the analysis (parameter = 1) or not (parameter = 0).
- Factor for multiplying species-specific widths of the dispersal kernel (α -values, the second column in your biodiversity feature list file). This parameter helps you to produce multiple solutions with different feature-specific scales of landscape use without needing to change the individual dispersal kernel widths manually after each run. All kernel widths will be multiplied with this factor.
- Value to indicate whether the program window will be left open (parameter = 0) or closed (parameter = 1). Closing the window at the end will allow the program to move on to the next run. If a batch file is composed for performing multiple runs, it is important to write 1 at the end of each line or the program will not proceed to the next Zonation run.

A batch file is run simply by double-clicking the file icon from windows. A batch file can also be used to run Zonation analyses through the graphical user interface (see section 4.).

A batch file is useful, for example, if you want to experiment with different levels of distribution smoothing, uncertainty analysis, BQP etc. (several settings files) or with alternative species weighting schemes (several species list files). You can also use batch files to run the most complicated analyses (requiring long computation times) overnight or over the weekend. When you wish to run multiple analyses, the simplest batch file consists of several command lines, each calling the program with different parameters, e.g.



```
call zig3 -r settings1.dat species.spp out\out1.txt 0.0 0 1.0 1
call zig3 -r settings2.dat species.spp out\out2.txt 0.0 0 1.0 1
call zig3 -r settings3.dat species.spp out\out3.txt 0.0 0 1.0 1
```

Here, the same species list file is run with three different settings. **Remember to identify the output files separately (give them different names), or the program will**

overwrite old results after each run.

Notice, however, that when performing multiple runs, you will by default only have the results that are automatically saved (see section 3.4.1. for file output). Remember to develop the settings for automated post-processing analyses if special processing of outputs are needed.

Additional command line parameters

In addition to mandatory regular command line parameters, there are optional parameters that do not affect the simulation results. These parameters can be in any order but they must always reside after the regular parameters.

```
--use-threads [n]
```

Use *n* simultaneous threads. If *n* is not specified, the program tries to use the number of hardware threads (e.g. number of CPUs or cores or hyperthreading units) available on the system. By default, Zonation uses a single thread. If neighborhood quality penalty (NQP) is set on, a single thread is used regardless of this option.

```
--image-output-formats [png] [bmp] [jpg] [emf]
```

With this parameter, you can generate output color map in the specified formats. Format specifiers can be in any order. No files are created if none of the formats are specified. By default, Zonation creates the color map in JPEG and EMF formats. The following files are created with each specifier ("output" is the specified output file without the filename extension):

`png`: output.png
`bmp`: output.bmp
`jpg`: output.jpg
`emf`: output.emf

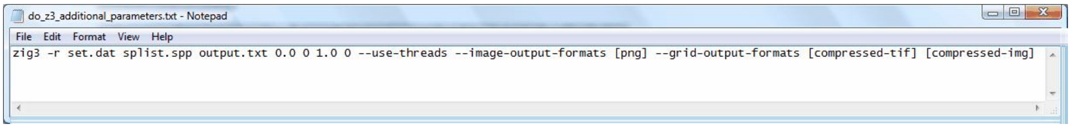
For backwards compatibility, EMF format produces exactly the same file as BMP. We do not recommend its use.

```
--grid-output-formats [asc] [tif] [img] [compressed-tif] [compressed-img]
```

With this parameter, you can generate Zonation output grids in the specified formats. This step will save you time if you want to work on the map further in a GIS. Format specifiers can be in any order. No files are created if none of the formats are specified. By default, Zonation creates the output grids in ASCII format - that you do not need to specify separately. The following files are created with each specifier ("output" is the specified output file without the filename extension):

`asc`: output.rank.asc output.prop.asc output.wrscr.asc
`tif`: output.rank.tif output.prop.tif output.wrscr.tif
`img`: output.rank.img output.prop.img output.wrscr.img
`compressed-tif`: output.rank.compressed.tif output.prop.compressed.tif output.wrscr.compressed.tif
`compressed-img`: output.rank.compressed.img output.prop.compressed.img output.wrscr.compressed.img

All output formats except ASCII have 32 bit float element type. Tif produces a GeoTIFF file and img an Erdas Imagine file. Compressed GeoTIFF uses DEFLATE compression.

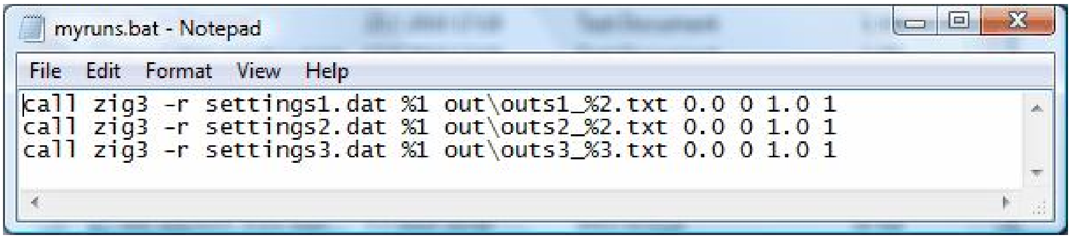

do_z3_additional_parameters.txt - Notepad
File Edit Format View Help

```
zig3 -r set.dat splist.spp output.txt 0.0 0 1.0 0 --use-threads --image-output-formats [png] --grid-output-formats [compressed-tif] [compressed-img]
```

An example of a batch file with additional parameters.

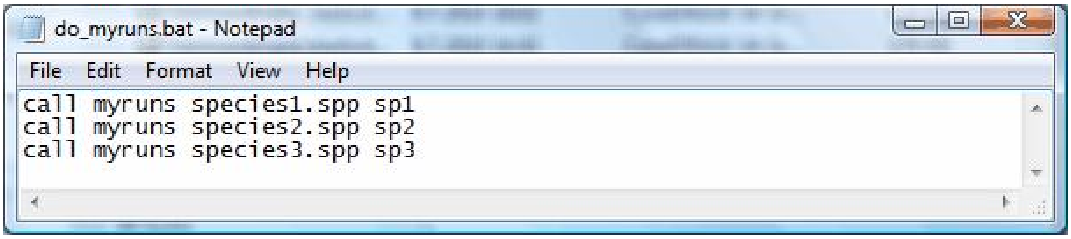
Advanced use of nested batch files

The description so far covers the basic use of batch files, but another more complicated example follows below. If you want to alter more than one parameter simultaneously, then one way forward is to use multiple nested batch files. For instance, you might want to run the above batch (do_zig.bat) with different species weights defined in separate species list files. To do this we first create a new file called "myruns.bat" and then adjust the settings from our example above:



```
call zig3 -r settings1.dat %1 out\outs1_%2.txt 0.0 0 1.0 1
call zig3 -r settings2.dat %1 out\outs2_%2.txt 0.0 0 1.0 1
call zig3 -r settings3.dat %1 out\outs3_%3.txt 0.0 0 1.0 1
```

Then we create another batch file, which in turn calls the "myruns"-batch.



```
call myruns species1.spp sp1
call myruns species2.spp sp2
call myruns species3.spp sp3
```

Here the first parameter after "myruns" defines which species list file is used when running the "myruns"-batch and the second parameter (sp1, sp2, sp3) defines a part of the output file name to distinguish which species file has been used in each run. In "myruns"-batch file these parameters will be referred to as %1 and %2. When running the "do_myruns"-batch the program first calculates solutions using "species1.spp"-file with three different settings and giving each of the output files an ending "sp1" (outS1_sp1, outS2_sp1 and outS3_sp1). Then the program repeats the procedure with the other two species list files, "species2.spp" and "species3.spp". Thus, running the batch produces nine solutions with different settings and species weights composition.

Using nested batch files is extremely useful when running many solutions using combinations of settings. For example, if we would use only one batch file to run the solutions described above, we would need to write a separate call for each of the nine solutions, and with many calls it is relatively easy to introduce errors into some of them.

Run simultaneous instances: zig3run

The **zig3run** utility can be used to run simultaneous Zonation instances on the command prompt. It is often more effective to use simultaneous processes than to use multiple threads in Zonation. On the other hand, multiple simultaneous single-threaded Zonation processes use more memory than a single process with multiple threads. The rule of thumb is that the more species layers you have, the more effective a multi-threaded Zonation becomes. With BQP set on, it is very efficient to use multiple threads.

Zig3run parses Windows batch files and tries to recognize calls to Zonation. It queues recognized calls and runs them simultaneously. It monitors the state of the running processes and prints a progress percent on screen. If a particular Zonation process is already running, it hooks up to the process and starts monitoring it until it finishes. A Zonation process is identified by the absolute path of its output file.

The Zonation calls in the batch file must each be on their own lines and they must not span multiple lines. After the call there must not be any other command. Zig3run will not recognize any constructs that affect the control flow in the batch files. For example, if you call a Zonation instance in a loop, only one call will be recognized.

Example batch files

A legal batch file

```
call zig3.exe -r set.dat splist.spp output.txt 0.0 0 1.0 0 --use-threads --image-output-formats
zig3.exe -r set2.dat splist.spp output.txt 0.0 0 1.0 0 --grid-output-formats asc tif
```

An illegal batch file

```
call zig3.exe -r set.dat splist.spp output.txt 0.0 0 1.0 0 \
        --use-threads
zig3.exe -r set.dat splist.spp output.txt 0.0 0 1.0 0 && zig3.exe -r set.dat splist.spp
output.txt 0.0 0 1.0 0
for %%f in (set.dat set2.dat) DO zig3.exe -r %%f splist.spp output.txt 0.0 0 1.0 0
```

Command line parameters

```
-h, --help
```

Print help

```
-i, --input-file arg
```

Input batch file

```
-m, --max-processes arg
```

Maximum number of simultaneous processes. Default is the number of hardware threads.

```
-u, --update-interval arg
```

Update interval of the process progress display in milliseconds. Default is 200.

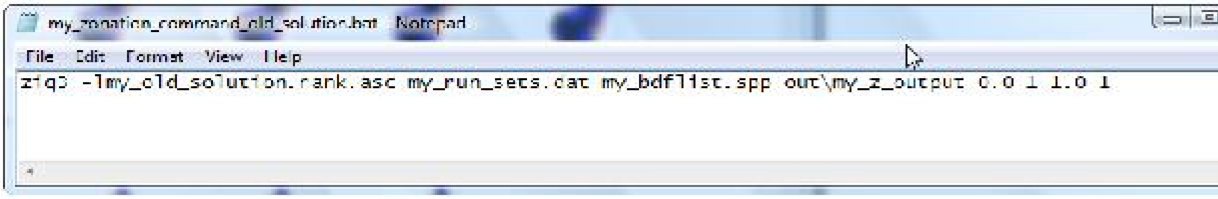
```
-o, --output-file arg
```

Appends all output from Zonation processes to a file. Useful for debugging.

3.2.2 Loading previously calculated Zonation solutions

It is also possible to load previously calculated Zonation solutions. This is a useful utility, if you want to make some further analysis with your old solution, but also if you need to test the performance of your old solution in different circumstances (see section 3.5.2. for solution cross-comparison using solution loading).

When operating the program from batch files, type "`-lfilename`" as the second parameter of your call. "`-l`" points out that an old solution is loaded. For file name enter the name of the priority rank file from your old solution (**.rank.asc file**, one of the output files produced during each run; see section 3.4.1.). Thus, a typical call when loading an older solution would look like this:


my_zonation_command_old_solution.bat - Notepad
File Edit Format View Help

```
zig3 -lmy_old_solution.rank.asc my_run_sets.dat my_bdflist.spp out\my_z_output 0.0 1 1.0 1
```

Remember to give a new output name in the call, if you do not wish to overwrite your old solution.

3.3 Input files & settings

3.3.1 Introduction

To use the Zonation software, you need a set of input files, some of which are compulsory, some optional. You can use some of the tutorial files as templates when creating your own input files.

When you enter file names, please note that **you need to specify the path to your file** as well. If your files are in the same folder with Zonation, the file name will suffice:

```
my_file.asc
```

However, most often it is convenient to have subfolders for different analyses. If the subfolders partition from the folder where you keep Zonation, please specify the path from that folder:

```
subfolder/my_file.asc
```

If you keep your Zonation and your analysis files in different places, e.g. Zonation in C:/Program files and analysis files at D:/Data, then you should specify the whole path to your analysis files:

```
D:/Data/subfolder/my_file.asc
```

This applies to all files that you refer to at any stage of your analysis setup. In the examples, we have left the paths out for the sake of simplicity.

3.3.2 Compulsory files

These three types of files are needed every time you run Zonation. There are many other optional files, but these three are compulsory.

(1) distribution grids of species or other biodiversity features, which describe the distribution and local density of each feature across the landscape (next section, 3.3.2.1.)

(2) biodiversity feature list file (section 3.3.2.2.) that indicates which biodiversity feature grids are to be used in the analysis. Certain settings, such as weighting, are adjusted in the biodiversity feature list file.

(3) run settings file (section 3.3.2.3.) that defines the settings and Zonation features to be used in your analysis.

3.3.2.1 Biodiversity feature map files

A GIS raster grid file of feature (species, habitat type, etc.) distribution, one file for each biodiversity feature. Note that in Zonation v.1.0. and v.2.0. these grids needed to be ascii grids (described below). Zonation v.3.0. uses the GDAL library for handling rasters, so in theory these grids can also be in any GIS raster grid format, whether ascii or binary. See the GDAL library for supported file formats. So far we have verified the function of .ascii .img and .tif formats, so we recommend using one of the three. Use other file types at your own risk. Unexpected errors and behaviour may occur when using the untested file types.

GIS raster files Zonation uses a variety of GIS raster files such as species distribution layers, cost layers, mask file or uncertainty map layer. All these files need to be exported from a GIS software (e.g. from ArcGIS) or produced with some other appropriate software. Zonation v.3.0 utilizes the GDAL library for the handling of GIS raster files. This adds flexibility compared to Zonation v.2.0, in which the species raster files and other rasters could only be in the ascii raster format. Now the raster grids can be entered in dozens of different formats (but see the warning above). Recommended formats include, for example, Erdas Imagine (.img). You can use the file name without a suffix wherever a grid raster can be used. Zonation will recognize the format automatically. The benefit of the alternative formats is that some of them save 80-90% of disk space compared to ascii rasters. They also load in only 10-20% of the time needed before. The only disadvantage is that only an ascii raster can be examined with a text editor such as notepad, but most of the other (binary coded) formats can only be examined using a GIS software. You can still use ascii rasters equally well in the v.3.0, if you prefer.

Note: The ability to use non-ascii rasters does not apply to the following specially processed integer grids:

- Removal mask layer
- Planning unit layer
- ADMU description layer

Whatever the grid file formats, it is important that all species distribution rasters have the same grid size! **This means that in all files the number of columns and rows as well as the size of cells should be equal**. It is equally important to have **at least one row of no data on each edge of the raster matrix.** This is due to computational reasons and the lack of these no data -rows leads to a situation where the program automatically transforms the values on edge rows to missing data. This in turn may alter species distribution information, which possibility one should be aware of. Missing data in the species distribution rasters do not necessarily need to be congruent between all species, the program will run if cells marked as "no data" for one species has values for other species. However, if missing data is not aligned, there are implications for the use of the BQP.

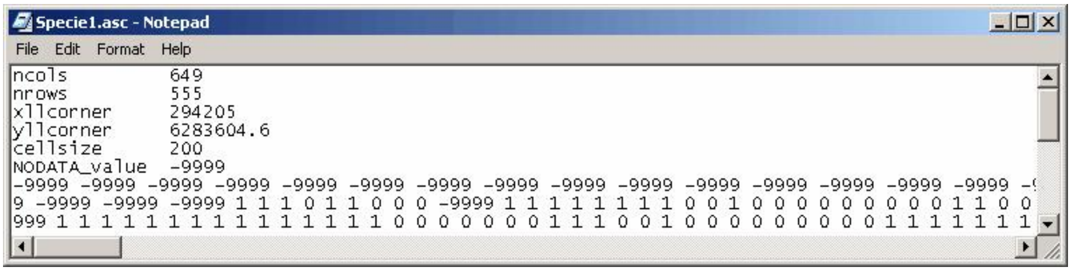
Description of ascii grid file format

Ascii files need to include the standard GIS raster file header

ncols	Number of columns
nrows	Number of rows
xllcorner	X-coordinates of the low-left corner
yllcorner	Y-coordinates of the low-left corner
cellsize	Cell size used in the raster file
NODATA_value	Definition of "no data" values. In example files "no data" has either value -1 or -9999.

After these rows comes the matrix showing the distribution of the species or other biodiversity feature. Each value in the matrix describes the occurrence of the feature in a specific cell. Values can be of any form of data, e.g. probability of occurrence, presence/absence -data, number of population etc. as long as the data is in same format across all biodiversity feature map files. Note that value 0 in the matrix indicates that the species does not occur in the cell with certainty, whereas lack of data must be marked as -1 or -9999, or with a similar value indicating "no data".

Remember also to use **decimal points, not commas**, in all the input files.



```
ncols         649
nrows         555
xllcorner     294205
yllcorner     6283604.6
cellsize      200
NODATA_value  -9999
-9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -
9 -9999 -9999 -9999 1 1 1 0 1 1 0 0 0 -9999 1 1 1 1 1 1 1 1 0 0 1 0 0 0 0 0 0 0 0 0 1 1 0 0
999 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 1 1 1 0 0 1 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1
```

Picture of biodiversity feature distribution map file, in this case with presence-absence information coded as 0 and 1.

In Zonation v.3.0., biodiversity feature map files can contain several types of biodiversity features and considerations. They include

- The distributions of individual species (observed or predicted occurrences, abundances or other measures of suitability)
- Distributions of community types (see section 5.3.1.)
- Negative conservation value of areas that are desirable for competing land uses (see section 5.3.4.)
- Distributions for the same sets of species, for multiple time steps (see sections 5.3.6. and 5.3.7.)
- Any properties of the habitats that can be assigned a value that reflects the conservation value of the site.

Different types of biodiversity features can be combined in a single analysis simply by listing them in the same biodiversity feature list file. In such an analysis, it is advisable to pay special attention to the values assigned to the features, as well as their weighting, to attain a desirable balance of prioritizing different kinds of features.

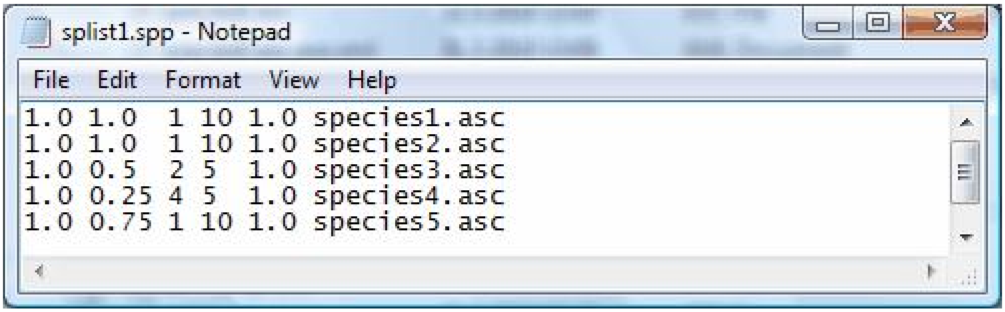
Please note also that Zonation can make use of point occurrence data as well as raster grids. Use of point occurrences is most appropriate when a species has only a few occurrences that are not enough to fit reliable distribution models to produce a comprehensive distribution map. The input for a SSI species is a (probably relatively short) list of observation locations instead of a map. See section 3.3.3.1. for details about using point occurrence data.

3.3.2.2 Biodiversity feature list file

A .spp file containing a list of all biodiversity feature map files, which will be used in your analysis. [Note that "species" is here used as a shorthand for any biodiversity feature, including species, habitat types, ecological communities, environmental conditions, ecosystem services, or any other information that acts as the basis for your decision making.].

.spp files A suffix for species list files such as biodiversity feature list file or uncertainty analysis weights file. In Zonation, these files would contain a list of species or other biodiversity features raster files with set-up parameters. You need to create all .spp files yourself, for example, with Notepad. You can use the tutorial files as templates. When saving a .spp file, remember to add the suffix (.spp) after the file name. You could also name these files .spp.txt to emphasize they are text files containing species lists.

In this file, each species is represented by one row of information. This file always contains six columns, except when using generalized benefit function as a cell removal rule, in which case there are nine columns (only in Zonation versions 2 and 3). The six-column version used by Zonation v1, v2 & v3 with the other cell removal rules has the following parameters:



```
1.0 1.0  1 10 1.0 species1.asc
1.0 1.0  1 10 1.0 species2.asc
1.0 0.5  2 5  1.0 species3.asc
1.0 0.25 4 5  1.0 species4.asc
1.0 0.75 1 10 1.0 species5.asc
```

Picture of biodiversity feature list file.

1. Species weight (w_j) in the Zonation algorithm. If no weights are used, this should be set to 1.0. In the example above all species have equal weights (1.0). Typically weights have positive values, but in Zonation v.3.0 weights can be zero or negative as well. Using 0.0 as a weight means that the program calculates the performance of the species during cell removal, but that the species does not in any way influence the priority ranking of the landscape. Thus, other species are acting as surrogates for weight-zero species. A feature with a negative weight is something undesirable that should be removed from the landscape as soon as possible. One way to think of negative-weight species is that they can represent (multiple) opportunity costs. Note that species-specific weights have no influence on the analysis if you have chosen the target-based benefit function as your cell removal rule. This is because all species have targets when using this cell removal rule - thus weights have no influence on the outcome.

   The weight parameter can include considerations such as degree of historical distribution loss, species local or global scale priority, taxonomic uniqueness, etc.

All such considerations should be aggregated to a single weight reflecting the importance of the species (biodiversity feature) in the analysis. Some notes about weighting: (i) Weighting is a political decision, and there is no general method for determining correct weights. (ii) Weights should still be used. If weights are implicitly taken as equal, then that is a weighting also. In analysis, a species with, say, unique taxonomic history, should have higher weight than a species that has 200 rather similar sister-species.

2. The α -value of species-specific scale of landscape use (parameter of negative exponential, α_j). **This parameter is only needed when you are using distribution smoothing or matrix connectivity as a part of your analyses, if not used, a dummy 1.0 will do fine**. The value indicates how species use the surrounding landscape and can be calculated based on, for example, the dispersal capability or the home range sizes of the species. The α –value can be calculated as

$$\alpha = \frac{2 * [\text{Cell size in km}]}{[\text{Use of landscape (km)}] * [\text{Input cell size}]}$$

E.g. if the known/guesstimated mean dispersal capability of a species A is 3 km and the cell size in species distribution files is 1.0 km (note, not 1000 m), then the value of α for species A is ~0.67.

$$\alpha = \frac{2*1}{3*1} \approx 0.667$$

The last part of the equation ([Cell size in km]/[Input cell size]) is needed to keep the α –**value in same unit of length as the cell size** given in the species distribution map file. E.g. if the cell size of previous example would be given in meters instead of kilometers (thus the cell size would be 1 000 m in the .asc files instead of 1 km), the α -value would be as follows:

$$\alpha = \frac{2*1}{3*1000} \approx 0.00067$$

Another common unit used in raster files is degrees. Also these need to be converted to get the correct α -value. Let us assume that the cell size in our example was 0.0083 degree, equaling approximately 0.860 kilometers. Thus the α -value in this case is

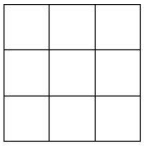
$$\alpha = \frac{2*0.860}{3*0.0083} \approx 67.1$$

It is important to understand, that this parameter is NOT the same α -value as is used in uncertainty analysis! The two parameters only happen to have been denoted with the same symbol in literature. Note also, that if distribution smoothing is not used, you should nevertheless enter a value in this column. This can be any positive number, e.g. a dummy value of 1. **Zonation will not run if this column is empty!**

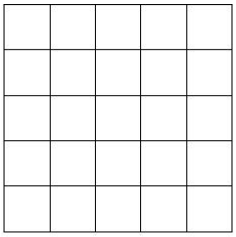
Columns 3 and 4 together define either BQP (Boundary Quality Penalty) or NQP (Neighborhood Quality Penalty in directed connectivity) settings for the species, depending on which one of the two options is used. The information of these two columns is only used if BQP or NQP is included to the analysis. Note, that even if BQP or NQP is not used, you should nevertheless enter a value into these columns. They can be any positive numbers, e.g. a dummy values of 1. **Do not leave the columns empty!**

3. When using BQP, this parameter indicates the row number in BQP definitions file linking the species to a correct penalty curve. When using NQP, this parameter serves the same purpose, but only for **upstream connectivity**. Thus, in NQP this number links the species to a penalty curve that describes how the value of focal planning unit changes when other planning units are lost upstream from the focal planning unit.

4. When using BQP, this parameter gives the species-specific buffer size (number of cells). The buffer size indicates the area around the focal cell in which any fragmentation (removal of cells) influences the quality of the focal cell. For species with large home ranges the buffer size should be larger and for species with small home ranges a smaller buffer size is adequate.



buffer size = 3



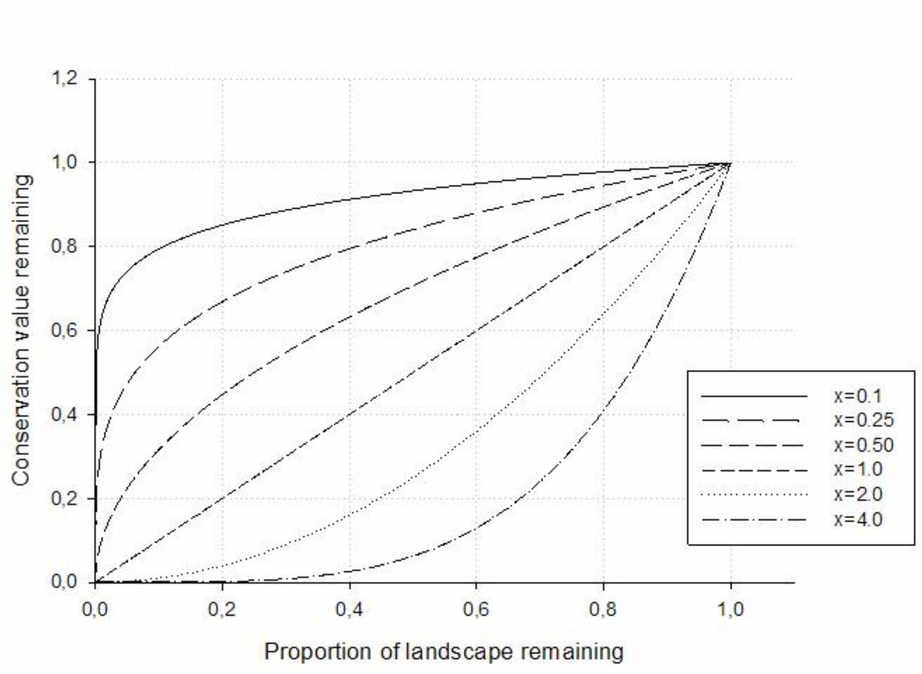
buffer size = 5

When using NQP, this parameter indicates the row number in BQP definitions file linking the species to a correct penalty curve, this time for **downstream connectivity**. Thus, when running NQP, every species have two penalty curves, one for upstream and another for downstream. As NQP-option uses planning units instead of singular cells, no buffers are needed. The connectivity of separate planning units, i.e. which ones are upstream or downstream, is defined in the directed connectivity layer.

Column 5 is used when using additive benefit function, target-based planning or generalized benefit function as your cell removal rule. With core-area Zonation a dummy number, e.g. 1.0 can be used.

5. This column has three functions depending on which cell removal rule is used.
   - If you are using **additive benefit function** as your cell removal rule, this parameter is the exponent (x) of the species-specific power function (r_{jx}) that

translates representation to value. The power function determines the rate of loss of conservation value from the remaining landscape as cells are removed. The exponent can be any positive number, but zero is not a valid value.



Picture of power functions with differing x -values.

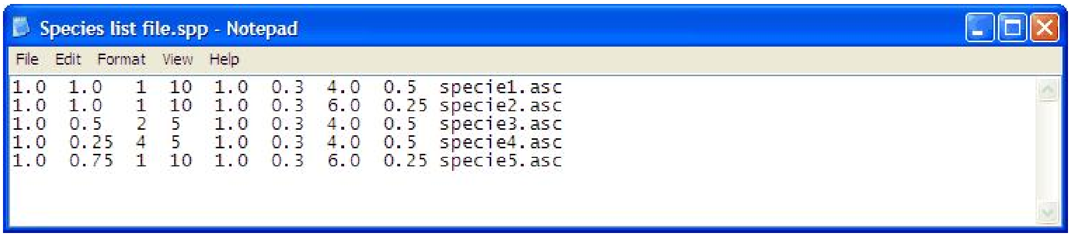
- If you are using **target-based planning** as your cell removal rule, this parameter determines the target proportion (from 0.0 to 1.0) of the species distribution, which you require in the final solution. *For a negatively weighted feature, this is the proportion of the feature that you want out from the solution (in Zonation v.3.0.).*

- If you are using **generalized benefit function** as your cell removal rule, then four extra parameters are needed for each species. Essentially, the final numerical column (the single ABF/TBF parameter) is split into four numerical columns, which in column order correspond to variables w_2, T_j, x, and y, in the two-piece power function. These parameters are used to determine the shape of the function, as explained in section 2.3.4. The fifth parameter, w_1, is the ordinary weight given for the species as the first column in the species list file.

Note that even though benefit functions or target-based planning is not used, you should nevertheless enter a value into the fifth column (however, no dummy values are needed to the three extra columns used with generalized benefit function). This can be any positive number, e.g. a dummy value of 1. **Do not leave the column empty!**

6. Name of the biodiversity feature map file (.asc -raster file). If your distribution

maps are in a different directory than your biodiversity feature list file, remember also to type the correct path in front of the name. Note that if you are using GBF as your cell removal rule, the contents of this column is shifted to column number nine.



```
1.0  1.0   1  10  1.0  0.3  4.0  0.5  specie1.asc
1.0  1.0   1  10  1.0  0.3  6.0  0.25 specie2.asc
1.0  0.5   2  5   1.0  0.3  4.0  0.5  specie3.asc
1.0  0.25  4  5   1.0  0.3  4.0  0.5  specie4.asc
1.0  0.75  1  10  1.0  0.3  6.0  0.25 specie5.asc
```

Picture of species list file when using generalized benefit function as cell removal rule.

There should not be any empty rows at the end of the species list file. If necessary, you can enter comments in your species list file on separate rows starting with the symbol #. Remember also to use **decimal points, not commas**, in all the input files.

[Supplementary material] Tips for using the command FOR to automatically produce species list files

Note that the `FOR`-command can be used in creative ways to automatically create species list files. For example, the following single command row, typed and run from the command prompt (or run from a .bat batch file),

```
FOR /L %i IN (1,1,900) DO @ECHO 1.0  1  1  1  0.25  p%i.asc >>
my_spp_list.spp
```

generates a file my_spp_list.spp, which has rows and relevant parameters for files p1.asc, p2.asc, ..., p900.asc. The ">>" at the end of the command indicates redirection of output into the following file. Without the ">> my_spp_list.spp", output is shown on the screen (command prompt).

Another variant of the for command allows one to loop through a set of files using the normal wildcard file name specification:

```
FOR %i IN (species_*.asc) DO @ECHO 1.0  1  1  1  0.5  %i
```

See the help for the `FOR`-command for further information.

3.3.2.3 Run settings file

A .dat file containing all basic Zonation settings. **This input file is needed only when running the program from the command prompt or batch files**. In the run settings file the following parameters need to be written on separate rows, with parameter names typed exactly as in the examples (case sensitive). If a parameter is missing from the run settings file, it will be assigned a default value.

.dat files	This is the suffix used in the tutorial and examples for the run settings file. Also this file needs to be created by yourself. You can use the tutorial files as templates. These files are also technically ascii files, that can be created using any text editor, including notepad.

It is very important that all the parameters in your run settings file are written **exactly** like presented here. If there are errors in the spelling of parameters, the program can not find them and will use default settings instead. The order of parameters on the other hand is not obligatory. If necessary, you can enter comments in your species list file on separate rows starting with the symbol # . Remember also to use **decimal points, not commas,** in all the input files.

The new settings allowed by Zonation v.3.0 (in addition to those in v.2.0) are:

Groups settings
```
use groups = 1
groups file = group_list.txt
```

Landscape condition and retention
```
use condition layer = 1
condition file = condition_layers.txt
use retention layer = 1
retention file = retention_file_name
retention layers relative weight  = β
```

Community analysis settings and matrix connectivity
```
[Community analysis settings]
load similarity matrix = 1
community similarity matrix file = comm_sim_matrix.txt
apply to representation = 1 OR

[Community analysis settings]
load similarity matrix = 1
connectivity similarity matrix file = matrix_file_name.txt
apply to connectivity = 1
```

Edge correction in connectivity
```
connectivity edge effect fix file = fname.asc
```

Administrative units
```
[Administrative units]
use ADMUs = 1
ADMU descriptions file = ADMU_weights_all_local.txt
ADMU layer file = fake_HV_ADMUs.asc
ADMU weight matrix = ADMU_weights_matrix.txt
calculate local weights from condition = 0
ADMU mode = 1
```

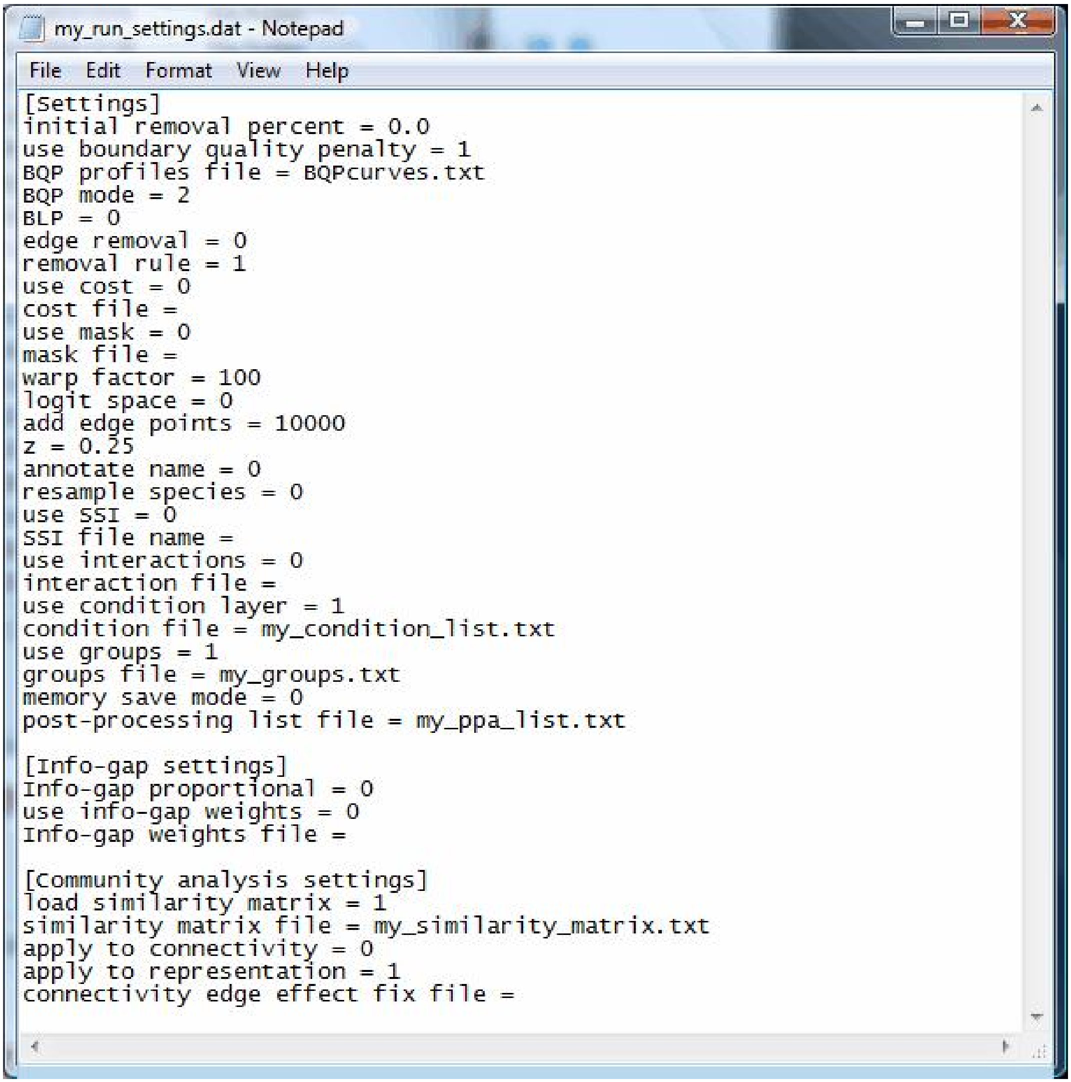
```
Mode 2 global weight = 1.0
```

Analysis area mask

```
mask missing areas = 1
area mask file = file_name
```

Automated post-processing

```
post-processing list file = ppa_list_file_name.txt
```



```
[Settings]
initial removal percent = 0.0
use boundary quality penalty = 1
BQP profiles file = BQPcurves.txt
BQP mode = 2
BLP = 0
edge removal = 0
removal rule = 1
use cost = 0
cost file =
use mask = 0
mask file =
warp factor = 100
logit space = 0
add edge points = 10000
z = 0.25
annotate name = 0
resample species = 0
use SSI = 0
SSI file name =
use interactions = 0
interaction file =
use condition layer = 1
condition file = my_condition_list.txt
use groups = 1
groups file = my_groups.txt
memory save mode = 0
post-processing list file = my_ppa_list.txt

[Info-gap settings]
Info-gap proportional = 0
use info-gap weights = 0
Info-gap weights file =

[Community analysis settings]
load similarity matrix = 1
similarity matrix file = my_similarity_matrix.txt
apply to connectivity = 0
apply to representation = 1
connectivity edge effect fix file =
```

An example of a settings file.

`[Settings]`

`removal rule`	Determines which cell removal rule (section 2.3.) will be used. 1 = Basic core-area Zonation, 2 = Additive benefit function, 3 = Target-based planning, 4 = Generalized benefit function, 5 = random removal. Default = 1.

`warp factor`	Defines how many cells are removed at a time. If warp factor is 100, it means that 100 cells are removed at each iteration. Thus a lower warp factor leads to a finer solution, but also to a prolonged running time, whereas a high warp factor keeps the running time short, but might result in a more coarse solution. If the warp factor is more than 1% of the remaining cells, then only 1% is removed. For example, if there are only 100 cells remaining in the landscape, then only one cell can be removed regardless of what the warp factor is. Note that if you are using planning units (PLU; section 3.3.3.11.) the warp factor is automatically set to 1. If planning units are not used, warp factor can be defined freely. [In our tests having a warp factor of 100 has had little influence on the solution compared to lower warp factor values, but the run times have been considerably shorter. We recommend to use a warp factor of 1 mainly for the final runs, if run times allow it. You can compare the effects of different warp factors with Landscape comparison.] Default = 1.
`edge removal`	Determines whether the program removes cells from the edges of remaining landscape (value = 1) or anywhere from the landscape (value = 0). Note that setting this parameter to 0 will increase the running times with large landscapes. Default = 1.
`add edge points`	Randomly selects additional cells inside the landscape that will be initially classified as edge-cells, from which removal can proceed. The value of this parameter determines the number of cells that are selected. This parameter allows a compromise between using and not using edge removal. Default = 0.

When adjusting the settings, it is good to understand the function of **Edge removal** and **Add edge points** options. The main profits of using edge removal is that it keeps computation times short and, to some extent, increases the connectivity of high quality habitats in the landscape structure. Hypothetically however, this option can have downsides in cases where a large area of "poor" habitat is completely surrounded by "good" habitats, and the Zonation program should first remove all the good habitats from the edge to reach the poor area. Naturally, by not selecting the edge removal option the program would easily find all poor habitats deep inside the landscape, but with the cost of lost structural connectivity and increased computation times. **To prevent any valuable areas to be lost and to keep the computation times short, using the add edge points option together with edge removal is recommended.** By adding hypothetical edge cells into the landscape the program can spot any larger poor areas surrounded by good habitats without the risk of removing valuable cells. You can of course mask out cities and other poor quality areas to provide Zonation sufficient low-quality edge from which ranking can naturally proceed.

`use SSI`	Determines whether Species of Special Interest (SSI) are included into the analysis (value = 1) or not (value = 0). The distribution data of these species is not given as maps but rather as a list of single point occurrences. If there are no SSI species to be included to the analysis, this parameter should be set to 0. Default = 0.

`SSI file name`	Similar to the biodiversity feature list file, except indicates a file that contains the list of SSI species used in the analysis (see section 3.3.3.1.). Default is that SSI is not used.
`use planning unit layer`	Determines whether a planning unit layer (section 3.3.3.11.) is used (value = 1) or not (value = 0). With this option cells are grouped into defined planning units which are then removed as a whole during the landscape ranking process.
`planning unit layer file`	Indicates the planning unit layer file (integer grid) to be used.
`use cost`	Determines whether land costs are included in the analysis (value = 1). If no land costs are used, this parameter should be set to 0. Default = 0.
`cost file`	Indicates the land cost file (section 3.3.3.6.) to be used. Default is that a cost file is not used. Note that this cost layer is used in cost-efficiency analysis where the conservation value of a cell is divided by local cost. If multiple opportunity costs are to be considered, these should be handled as negatively weighted grid layers (see section 5.3.4.).
`use mask`	Determines whether a removal mask layer (section 3.3.3.9.) is used (value = 1) or not (value = 0). Default = 0.
`mask file`	Indicates the mask layer file (integer grid) to be used. Default is that a mask file is not used.
`use boundary quality penalty`	Determines whether BQP connectivity is used (value = 1) or not (value = 0). Use of BQP leads to solutions that include structural aggregation at scales relevant for individual species (features). Default = 0.
`BQP profiles file`	Indicates the BQP profiles file (section 3.3.3.2.) to be used. Default is that a BQP profiles file is not used.
`BQP mode`	Determines how the program will calculate the effects of fragmentation from species distribution data. Essentially, this parameter tells the program what type of species distribution layers you are using. The mode can be either 1 or 2 depending on your data (see section 3.3.3.2.). Default = 2.
`BLP`	Defines a penalty given for the boundary length of the reserve. The value of BLP should be a small decimal number. The value of this penalty cannot be decided in advance based on some numeric criterion; rather a suitable value needs to be found by experimentation. It is desirable that only small loss of

representation follows from the use of BLP. Use of BLP leads to a more aggregated solution. Try first a small value, e.g. 0.01, to perceive the effect of BLP to the solution. **When including BLP in the analysis, preferably use a warp factor of 1**. If BLP is not used, this parameter should be set to 0. Note potential difficulties in interpretation of results if multiple aggregation methods are used simultaneously. Default = 0.

`use tree connectivity`	Determines whether directed connectivity (Neighborhood Quality Penalty; NQP) is used (value = 1) or not (value = 0). Unlike the other aggregation methods, NQP models a directed connectivity measurement up and down a linked system of planning units, which would typically model hydrologically linked water catchments. Note that planning units need to be always used together with NQP. Default = 0.
`tree connectivity file`	Indicates which tree connectivity file (section 3.3.3.3.) will be used when including NQP into the analysis. Default is that the file is not used.
`use interactions`	Determines whether ecological interactions (section 2.6.) are included (value = 1) into the analysis or not (value = 0).
`interaction file`	Indicates which interactions definitions file (section 3.3.3.8.) will be used.
`annotate name`	With this option you can mark your output file names to show which analyses have been used to produce them (value = 1). The program will add letters and numbers in the middle of your output file name depending on the used analyses: CAZ_/ABF_/TBF_/GBF= shows, whether you have used the basic core-area Zonation (CAZ), the additive benefit function (ABF), the target-based function (TBF), or generalized benefit function (GBF) as your cell removal rule. M = mask used. C = costs used. E = edge removal used. A = edge points added. Sxxx = distribution smoothing used. The following numbers show the factor that has been used to multiply the species-specific α –values. Note that for output the factor has been multiplied by 100. Thus using factor = 1.0 would result a suffix of S100, factor = 0.1 results a suffix of S10 etc. IGxxx = uncertainty analysis included. Again the following numbers show the info-gap α –value multiplied by 100. BQP = BQP included. BLPxxxx = BLP included. The following numbers show the penalty given for the boundary length multiplied by 1 000. I.e. using BLP = 0.5 results a suffix of BLP500,

	BLP = 0.05 results a suffix of BLP50 etc. Default = 1.
`logit space`	A special option relevant for probability of occurrence models using logistic link functions. Determines whether the biological values of cells will be transformed from logit space (value = 1) for processing. In this case the raster files (.asc files) should contain the values of the linear predictor part of a logistic link function. If data is not to be transformed from logit space, this parameter should be set as 0. Default = 0.
`treat zero-areas as missing data`	This option changes all cells with no species occurrences to missing data. This function might be useful in some cases, for example, if the missing data is in fact marked with the value 0 in your species distribution files due to some technical reasons. Note however, that there is a fundamental biological difference between species not occurring somewhere (value 0) and us not having any information from that same place (missing data). Thus, use this option with care. The use of this option does not change your input files in any way (thus cells with value 0 will remain as they are), it will only change the way Zonation interprets the files. Default = 0.
`z`	This value is used to calculate the extinction risks of species as their distribution sizes are decreasing. z is the exponent of the species-area curve ($S = cA^z$), which has been widely used in ecological studies. In theory, you can give z any positive value, but a commonly found empirical value is 0.25. If using an exponent equal to *z* in an ABF analysis, then Zonation is essentially minimizing the SA curve predicted extinction risk across species. Default = 0.25.
`resample species`	This option allows you to, for example, test analyses using only a subset of species. The program selects a random set of species from your species list file and uses them to run the analysis. Thus you can run several analyses and check how the selection of species influences the outcome. Note that the random sets do not include multiple selections of one species, all species in the set are different ones. To use this option enter the number species you wish to include in one set. If this value is zero, any negative value or equal to the total number of species, no sampling is done. Default = 0.
`post-processing list file`	Indicates the list file of automatically executed post-processing analyses (sections 3.3.3.17. and 3.5.1.) to be performed after landscape ranking. Default is that automated post-processing is not used.
`memory save mode`	Here you can choose to use Zonation in a memory save mode (value = 1). This can be useful if your analysis is so big that the physical memory of your computer just runs out. In other words,

	using this option allows slightly bigger analyses to be done. [Best solution for memory problems is computer with more memory.] If this value is set to zero, Zonation will not operate in the memory saving mode. Default = 0.
`use groups`	Determines whether a groups file (section 3.3.3.12.) is used (value = 1) or not (value = 0). A groups file is needed for grouped output, and also when condition or retention is to be used. Default = 0.
`groups file`	Indicates the groups file to be used in the analysis. Default is that the file is not used.
`use condition layer`	Determines whether feature-group specific landscape condition (section 2.10.) is used in the analysis (value = 1) or not (value = 0). Default = 0.
`condition file`	Indicates the file describing linkage of features to landscape condition (section 3.3.3.14.).
`use retention layer`	Determines whether feature-group specific retention layers (section 2.10.) are used in the analysis (value = 1) or not (value = 0). Default = 0.
`retention file`	Indicates the file describing linkage of features to landscape retention (section 3.3.3.15.). Default is that retention file is not used.
`retention layers relative weight`	Determines the relative weight of retention layers (as a group) compared to representation (normal grids not transformed for retention). Default = 1.0.
`mask missing areas`	Determines whether some areas of the landscape are masked out (filled with missing data) for all input grid layers (value = 1) or not (value = 0). This option is useful if only a subregion of the landscape would be analysed and one does not wish to redo all input grids. Default = 0.
`area mask file`	Indicates the raster file to be used for masking the areas with missing information (section 3.3.3.10.). Default is that area mask file is not used.

`[Info-gap settings]` This title in brackets is obligatory before the info-gap settings.

These settings are for including uncertainty in distributions (section 2.5.) into your analysis.

`Info-gap proportional`	Determines whether the errors in species occurrences are uniform errors (value = 0) or proportional errors (value = 1). Uniform error is the default setting and works for most of the data sets, but in some cases it is more appropriate to use proportional errors (see e.g. Ben-Haim 2001).

`use info-gap weights`	Determines whether species-specific distribution uncertainty map layers (section 3.3.3.7.) are used in the info-gap analysis (value = 1) or not (value = 0). Default = 0.
`Info-gap weights file`	Indicates the file that includes the list of uncertainty maps (that correspond to species grids; section 3.3.3.7.). Default is that the file is not used.
`[Community analysis settings]`	This title in brackets is obligatory before the community analysis settings. Note: the capability to use the following community-level analyses is new in Zonation v.3.0.
`load similarity matrix`	Determines whether a similarity matrix for connectivity or community composition (section 3.3.3.4.) is used in the analysis (value = 1) or not (value = 0). Default = 0. If the parameter is set to 1, either similarity expansion for representation, or similarity in connectivity, or both, can be accounted for, as specified by the following parameters.
`connectivity similarity matrix file`	Indicates the connectivity similarity matrix file (section 3.3.3.4.) to be used. Default is that similarity matrix file is not used.
`apply to connectivity`	Determines whether the similarity matrix is applied to connectivity (section 2.4.5.) in conservation area prioritization (value = 1) or not (value = 0). Default = 0.
`connectivity edge effect fix file`	Indicates the file for edge effect fix for connectivity (section 3.3.3.5.). Default is that edge fix file is not used.
`community similarity matrix file`	Indicates the community similarity matrix file (section 3.3.3.4.) to be used. Default is that similarity matrix file is not used.
`apply to representation`	Determines whether community similarity expansion (section 2.8.) is applied to the stack of biodiversity feature grids. (value = 1) or not (value = 0). Default = 0.
`[Administrative units]`	This title in brackets is obligatory before the administrative units settings. Note: the capability to use the following administrative units settings is new in Zonation v.3.0.
`use ADMUs`	Determines whether administrative units (section 2.12.) are accounted for in the analysis (value = 1) or not (value =0). Default = 0. If 1, all information below needs to be given for analysis to succeed.
`ADMU mode`	Determines whether the administrative regions are weak (value

	= 1) or strong (value = 2) See section 2.12. for details. Default = 1.
`ADMU layer file`	Indicates the administrative units layer (integer grid; section 3.3.3.16.) to be used. Compulsory if ADMUs are used.
`ADMU descriptions file`	Indicates the file describing weights etc. for the administrative units (section 3.3.3.16.). Compulsory if ADMUs are used.
`ADMU weight matrix`	Indicates the file containing a matrix of administrative units x feature weight (section 3.3.3.16.). Compulsory if ADMUs are used.
`Mode 2 global weight`	Indicates balance between global representation and local (ADMU-specific) considerations in landscape ranking. When ADMU mode 2 is used, this is the weight given to global weights of species or other biodiversity features, from the perspective of the global decision maker. The value can vary between 0 and 1. Default = 0.5.

3.3.3 Optional files

These files are needed only when certain options are used.

3.3.3.1 SSI list and coordinates

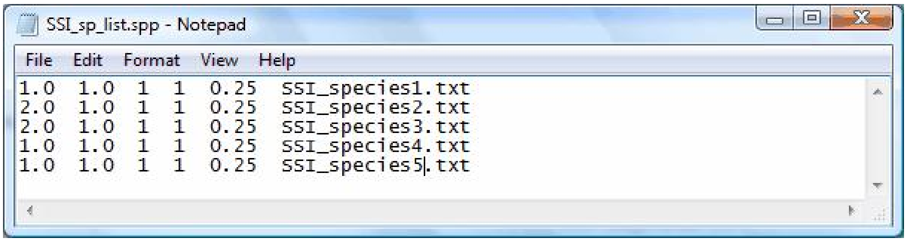
The SSI species (Species of Special Interest) are the second kind of species occurrence information that can be entered into Zonation. The input for an SSI species is a (probably relatively short) list of observation locations instead of a map. The SSI species input can be used for a species that either (i) has so few observations that the distribution of the species cannot be modeled, or (ii) has been completely surveyed and all occurrence locations are known. The idea in the SSI species is that if the species only occurs in a few locations, it is wasteful to enter a full map for it; a one million element grid map takes ~3000 times as much memory to store as does a list of 100 observation locations + population sizes. Consequently, a very high number of SSI species can be analyzed in Zonation. Ordinary 'map' species and SSI species can be mixed in the same analysis, however, it is not currently possible to run Zonation only with SSI species. (But one can just enter one zero-weighted map of the landscape and all of the rest of the species as point distributions, which amount to an observed-distribution-only analysis.)

SSI species are treated exactly as map species in the Zonation process, the marginal loss following the removal of a cell is based on the fraction of the distribution of the species residing in the cell. However, there is the difference that distribution smoothing and boundary quality penalty connectivity methods do not operate on SSI species. Note that connectivity requirements for an SSI species can be implemented indirectly by entering "buffering" locations for the SSI species around the actual occurrence locations. Overall, it can be expected that full distributions of SSI species will be retained far into the cell removal process especially if there are relatively few locations with observations of these species.

To include SSI species into the analyses, you need two kinds of input files:

1) SSI species list file
2) Species-specific coordinate files

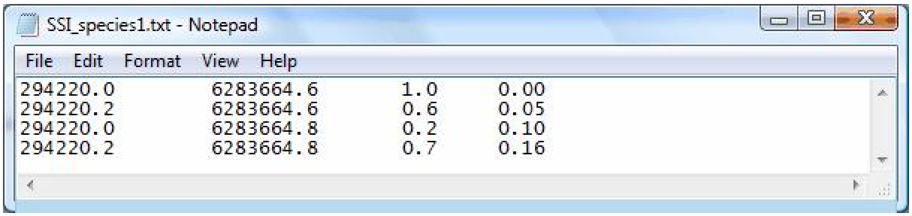
SSI species list file has an identical structure to the ordinary species list file. However, it should be understood that the columns for the dispersal alpha and BQP parameters contain dummy values – whatever is entered there will not influence computations. Thus, the relevant columns for an SSI species are (i) the first one defining the weight of the SSI species, (ii) the last numeric column which gives the parameter either for additive benefit function or for targeting analysis and (iii) the last column indicating the coordinate files.



```
1.0  1.0  1  1  0.25  SSI_species1.txt
2.0  1.0  1  1  0.25  SSI_species2.txt
2.0  1.0  1  1  0.25  SSI_species3.txt
1.0  1.0  1  1  0.25  SSI_species4.txt
1.0  1.0  1  1  0.25  SSI_species5.txt
```

Picture of SSI species list file.

The second set of files are the species-specific coordinate files, one for each species, which give the exact locations for each observation. An SSI species distribution file has four columns: 1) x coordinate of the observation point, 2) y coordinate of the observation point, 3) the biological value of the record (this can be population size or other measure of site suitability for the species in question) and 4) the info-gap relative error measure, with exactly the same interpretation and operation as for the map species.The given coordinates must fall within the area of the maps loaded for map species (as defined in the species distribution map files for ordinary map species) or otherwise an error will be reported. If uncertainty analysis is used, distribution discounting will be applied to the population size (or to any other form of information) given for the location. Note that the fourth column can be omitted for SSI species; if so, the uncertainty error measure will be taken as zero, and any uncertainty analysis will not influence population sizes given in column number three.



```
294220.0        6283664.6       1.0     0.00
294220.2        6283664.6       0.6     0.05
294220.0        6283664.8       0.2     0.10
294220.2        6283664.8       0.7     0.16
```

Picture of the species-specific coordinate file. The first two columns give the x and y coordinates of the record. The third column shows the biological value of that record (any non-negative integer or decimal value) and the last column is the relative error measure.

Run settings for using SSI species

When including SSI species into your analysis, remember to type in to your Run settings -file "`use SSI = 1`" (SSI option selected) and "`SSI file name = my_SSI_list.txt`" (name of your SSI species list file).

Output with SSI species

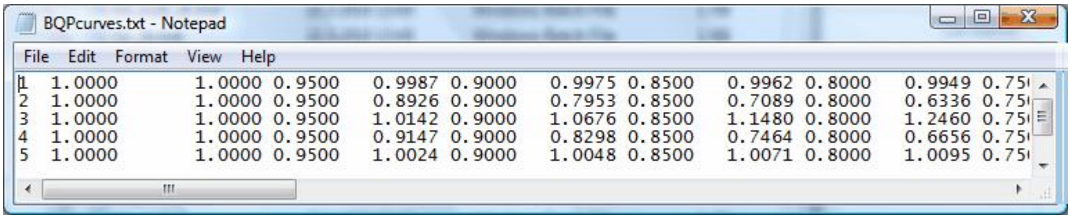
Numbers for the mean and minimum representation curves are given in a special .SSI_curves.txt output file, which is produced together with the basic output files whenever SSI species are included into the analysis. In this file there is also information about the level of landscape removal when the last occurrence of that SSI species is removed. Output for SSI species is given in the "Plots" tab in the GUI Runtime viewer, where a graph displays the minimum and mean fraction retained across all SSI species. Locations with SSI observations are shown by red in the "Maps" -window – it is worth checking that the SSI locations display correctly as errors in coordinates might otherwise easily go unnoticed.

3.3.3.2 Boundary quality penalty definitions file

To include boundary quality penalty (BQP; for a description of the method see section 2.4.3.) in to your analysis you need to:

1. Create a BQP definition file which contains all penalty curves. This file determines different responses of species to habitat fragmentation.
2. Link all species to the correct penalty curve by entering the correct row number of the respective curve into the third column of your biodiversity feature list file. Multiple species can, and commonly will, link to the same response curve.
3. Give a suitable buffer size (in cells!) for each species in your biodiversity feature list file. The buffer size indicates the area in which any habitat loss and fragmentation will influence the biological value of the focal cell for that particular species.
4. Decide how you want Zonation to treat missing data in a BQP analysis.

BQP definition file is a text file where different species responses to neighborhood habitat loss are displayed as points of penalty curves, each curve on their own row.



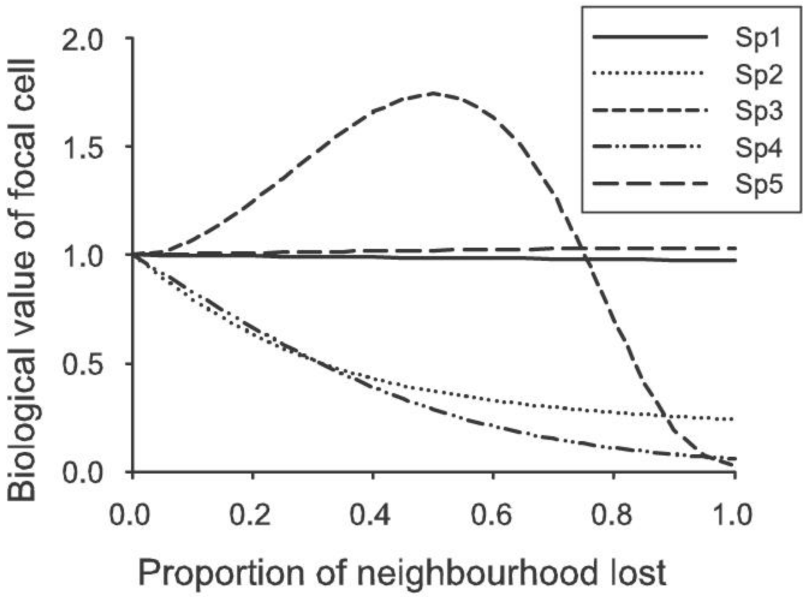
```
1  1.0000        1.0000 0.9500   0.9987 0.9000   0.9975 0.8500   0.9962 0.8000   0.9949 0.75
2  1.0000        1.0000 0.9500   0.8926 0.9000   0.7953 0.8500   0.7089 0.8000   0.6336 0.75
3  1.0000        1.0000 0.9500   1.0142 0.9000   1.0676 0.8500   1.1480 0.8000   1.2460 0.75
4  1.0000        1.0000 0.9500   0.9147 0.9000   0.8298 0.8500   0.7464 0.8000   0.6656 0.75
5  1.0000        1.0000 0.9500   1.0024 0.9000   1.0048 0.8500   1.0071 0.8000   1.0095 0.75
```

Picture of BQP definition file.

Here, the first column indicates the number of row. After row numbers comes the first column pair, in which the initial state (when no neighborhood habitat has been lost) is represented. The first number of the pair indicates the proportion of neighborhood habitat left and the second column indicates the fraction of biological value retained in the focal cell. Hence, in the initial state, when no habitat has yet been lost and therefore the biological value of the focal cell has not yet changed, the two parameters always have a value of 1.000, respectively. The following column pairs describe the loss of neighborhood habitats and effects that this has on the biological value in the focal cell. Note that the change in biological value can be either negative (retained < 1) or positive (retained > 1) depending on species preference to fragmented habitats.

As mentioned earlier, the column pairs should be considered as x,y -coordinates on a penalty curve. To draw a penalty curve, or any curve at all, you need to have at least two points. Thus, in the BQP definition file at least two column pairs are needed. The two points could for example be the initial point, when no habitat has been lost and the final point, when all the habitat has been lost. E.g. for species A the two points could be (1.000 1.000) and (0.000 0.500), meaning that when all the neighborhood habitat has been lost, the biological value of the focal cell for the species has decreased by half. Note that each of the penalty curve rows can contain a maximum of 20 points!

These BQP functions can be defined based on statistical analysis of habitat models, or based on expert knowledge; see Moilanen and Wintle (2007).

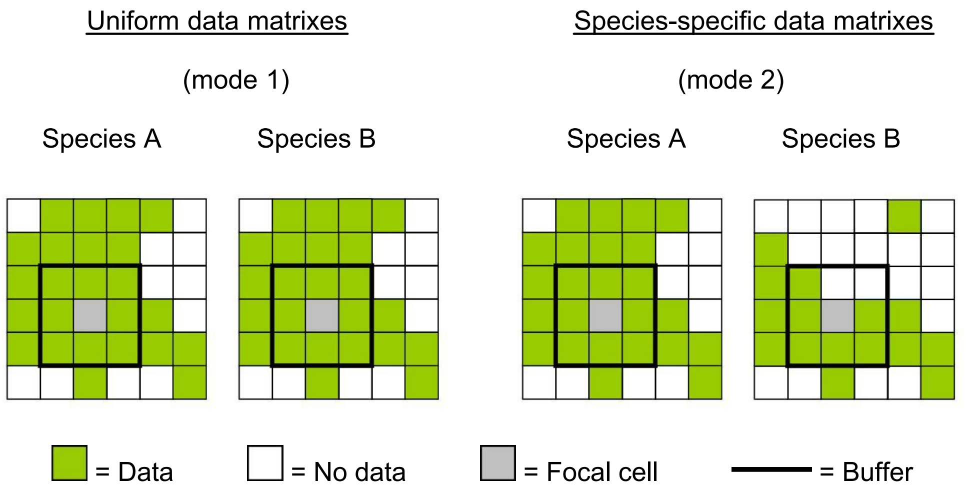


The figure above represents an example of different species-specific penalty curves (redrawn from Moilanen & Wintle 2007). Note that the curve increasing over value 1 indicates a species that prefers semi-fragmented habitats.

Boundary quality penalty can be included in two alternative modes. Your choice will depend on your species (or other biodiversity features) distribution grids. BQP mode is defined in the run settings file.

Mode 1 indicates that the data/no data matrix in all species distribution map files should be uniform and aligned, and that there are no differences between species in terms of which cells are considered potential habitat, and which are then used in BQP buffer calculations. In other words, all species would be dependent on the same general habitat type, such as forest. With mode 1, Zonation automatically aligns missing data if different species layers happen to have missing data at different locations. When aligning data, if species A has missing data at location (x,y) where any species B has a positive occurrence, then the missing value for species A is replaced by a zero-level occurrence.

Mode 2 indicates that the data/no data matrixes are not uniform and aligned, and that the program needs to calculate species-specific buffers for each species/cell separately. Mode 2 is more realistic in the sense that fragmentation/loss in habitats that are not suitable for the species will not influence the value of the focal cell. But, mode 2 also requires longer computation times due to more complicated species-specific calculations. Also, use of mode 2 at least doubles the memory usage of Zonation, thus decreasing the number of species that can be run in one analysis. Thus, mode 1 is a preferable when all species use approximately the same habitat type.



Run settings for using BQP in your analysis

To run BQP from command line, you need following lines in your run settings file:

```
use boundary quality penalty = 1
```

(option selected)

```
BQP profiles file = myBQPdefinifionsfile.txt
```

(the name of your BQP definitions file)

```
BQP mode = 1 OR 2
```

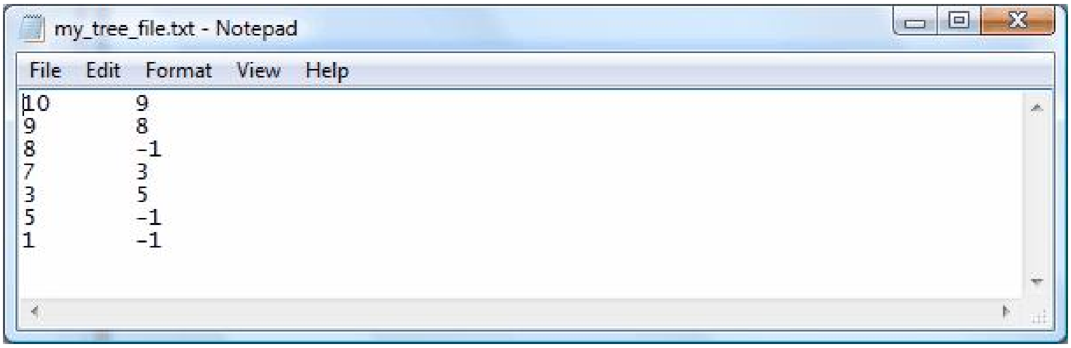
(depending on how you want Zonation to treat missing data about species. See BQP mode in section 3.3.2.3. for details.)

3.3.3.3 Directed connectivity layer

To include directed connectivity (NQP) to your analysis you need to:

1. Give a planning unit layer (section 3.3.3.11.) to identify which cell belongs to which planning unit.
2. Create a directed connectivity definitions file describing the linkage between planning units.
3. Give both **upstream and downstream** connectivity responses for all species. These responses are defined in a file that is identical to a BQP definition file (section 3.3.3.2.) with the distinction that instead of one, every species is linked to two penalty curves.
4. Link all species to the correct penalty curves by entering the correct row number of the respective curve into the **third (upstream) and fourth (downstream) column** of your biodiversity feature list file.

Directed connectivity definitions file is a text file that contains a description of the tree hierarchy, i.e., the linkages between planning units (PLU). This information is entered simply as file with two columns, a planning unit number (as given in your planning unit layer) and the number of the planning unit downstream, as in the figure below.



```
10        9
9         8
8         -1
7         3
3         5
5         -1
1         -1
```

This would be interpreted so that the PLU 10 flows into PLU 9, PLU 9 flows into PLU 8, and PLU 8 has no downstream component – for example, if it flows into the sea. PLU 7 flows into PLU 3, 3 into 5, and 5 is at root of the tree. PLU 1 is unlinked to anything and it is taken as an independent entity. The ending of a linkage line is always marked as -1. Note that the planning unit numbers need not be consecutive. Warnings will be issued to the memo if linkage information is missing for a planning unit, or linkage is confused, like when having multiple downstream connections for one planning unit (remember that a PLU can have several upstream connections, but only one downstream connection).

In addition to the directed connectivity layer, loss functions (i.e. penalty curves) completely analogous to those used in the BQP technique need to be defined for each species. Just like with BQP, the penalty curves represent the loss of biological value in the focal unit (here the planning unit) when neighboring units are removed. The difference to the BQP is essentially that the neighborhood is not symmetric, but directional, and that for each species there are separate upstream and downstream response functions modeling effects of habitat loss upstream or downstream from the focal location (hence, instead of one, each species may be linked to two penalty curves).

Note that the use of NQP also changes the interpretation of the biodiversity features list file. With the NQP, there is no species-specific radius like with BQP – the neighborhood is the set of linked planning units. Rather, instead of as number of response and radius, the third and fourth columns in the species list file are interpreted as the row number of the penalty curves for upstream and downstream losses, respectively. The curves are specified in the same input file as in BQP (see section 3.3.3.2.).

Run settings for using directed connectivity in your analysis

To use directed connectivity in your analysis from command line, you need to:

1. Set "`use planning unit layer`" to 1 in your run settings file
2. Give the name of your planning unit layer file "`planning unit layer file = my_PLUs.txt`"
3. Set "`use tree connectivity`" to 1 to indicate that NQP will be used.
4. Also define the name of your directed connectivity definitions file in the run settings file with "`tree connectivity file = mytreeconnectivity.txt`".
5. Define the name of your BQP curves file by typing "`BQP profiles file = myBQPdefinifionsfile.txt`"
6. Note that when planning units are used, the program will automatically set warp

factor to be one, regardless what has been defined in the run settings file.

3.3.3.4 Connectivity and community similarity matrices

Connectivity matrix

A connectivity matrix is a file that is used in an analysis to apply matrix connectivity. The file is a numeric matrix with N columns and N rows and describes the extent to which the occurrence level of a biodiversity feature influences the connectivity of multiple other biodiversity features (see section 2.4.5.). The values describe pair-wise connectivity effects, i.e. how much feature *n* (column) influences the connectivity of feature *k* (row). The matrix does not need to be symmetric; feature *n* may contribute more or less to the connectivity of feature *k* than the other way around. When Zonation reads in the connectivity matrix, it is applied to the first N features in the biodiversity feature list file.

Run settings for including connectivity matrix

To include the connectivity matrix in to your analysis, adjust your run settings file to include the following lines:

`[Community analysis settings]`
`load similarity matrix = 1`
`connectivity similarity matrix file = connectivity_matrix.txt` (the name of your connectivity matrix)
`apply to connectivity = 1`

Please refer to section 5.1.6. for more detailed analysis setups.

Community similarity matrix

The community similarity matrix is applied in the community level analyses. It describes the similarity between community types - most commonly this would be the proportion of species in common between two community types, but it can be other things as well. The values can vary between 0.0 and 1.0 and describe pair-wise similarities between community type A (in rows) and B (in columns). The values are used to expand the occurrence levels of the primary community types to effective occurrence levels, in the extent to which features of community type A are present in type B and all the other types. The matrix does not need to be symmetric; community type A may contain more features of type B than the other way around. The dimension of the matrix does not need to be identical to the number of biodiversity features used in the analysis. When Zonation reads in a community similarity matrix with N rows and N columns, it applies it to the first N features in the biodiversity feature list file. This allows for combined community and species level analysis (see section 5.3.2.)

A community similarity matrix is applied in a different way from connectivity matrix. Community similarity is used for transforming the actual occurrences of community types, whereas connectivity matrix is used to transform connectivity calculations. Distribution smoothing following community similarity expansion would have a similar effect to matrix connectivity.

Run settings for applying a community similarity matrix

To include the community similarity matrix in your analysis, adjust your run settings file to include the following lines:

```
[Community analysis settings]
load similarity matrix = 1
community similarity matrix file =
community_similarity_matrix.txt
apply to representation = 1
```

Please refer to sections 5.3.1. and 5.3.2. for more detailed analysis setups.

3.3.3.5 Connectivity edge effect fix file

Connectivity edge effect fix file can be applied in an analysis that utilizes matrix connectivity. It is a raster grid (.ascii or .img or .tif), in which values in cells indicate the fraction of the cell that belongs to a habitat that does not harm connectivity. This fix can be useful, for example,

(i) To account for national borders, beyond which suitable habitats may continue, but the connectivity seems to be lower on the edge, as no data is available from the other side. Cells on the other side of the national border could be marked as non-harmful base habitat. Then, it is effectively assumed that habitat outside the border will influence connectivity as habitat inside the border. If the connectivity of the cell is 2.0, which has been aggregated from a neighborhood that is only 1/3 within the country, then if cells outside the country have been marked as, the value of connectivity becomes 2.0/(1/3)=6.0.

(ii) In a situation where a mosaic of different habitat types is beneficial for biodiversity, but the distribution maps implicate that patches of different habitat types have discontinuous and patchy distributions. For example, thinking of connectivity of a forest. The connectivity of forest will necessarily be reduced at the edge of a large lake. But, there may be many cases where such an edge-effect is not desirable. If not, mark water as non-harmful base habitat in the connectivity edge effect fix file. Or, this option could be relevant at the border of a forest and a marshland (some species will not perceive the marshland as bad for connectivity).

Run settings to include the connectivity edge effect fix

To include connectivity edge effect fix for matrix connectivity, you first need to adjust your run settings to include matrix connectivity (see section 3.3.3.4.). In addition, you need the following line:

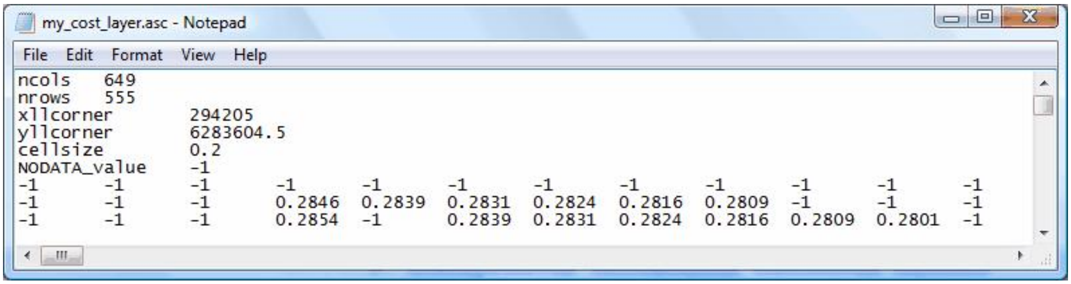
```
connectivity edge effect fix file = fixfile_name.asc
```

so that Zonation will also read in the edge effect fix file.

In addition, you may want to correct for the amount of habitat within each cell through the use of a cost layer (`use cost = 1`, and `cost file = habitatproportion_name.asc`). In that case the cost layer should include the proportion of habitat in each cell (see section 2.3.6.).

3.3.3.6 Cost layer

A standard GIS raster file (.ascii or .img or .tif) about land cost**.** This file includes all basic raster information as explained in species distribution map files and a map of land costs in each cell. The land cost value in the matrix can be any positive number larger than 0. Thus, zero nor any negative value can not be used as land cost! If areas with no land costs need to be included in the analyses, the land cost value for these cells can be set very low, e.g. 0.00001. Note that here the costs do not need to be measured in terms of money, you can use other measures of economical loss as well. For example in economical fisheries the fishing intensity of a landscape can be used as a cost layer - the higher the fishing intensity, the higher is the cost of protecting the particular site. The cost layer is an optional file, Zonation would most commonly be run without cost data. If land costs are not included in the analysis, all cells implicitly have an equal cost value of 1.



```
ncols   649
nrows   555
xllcorner       294205
yllcorner       6283604.5
cellsize        0.2
NODATA_value    -1
-1      -1      -1      -1      -1      -1      -1      -1      -1      -1      -1      -1
-1      -1      -1      0.2846  0.2839  0.2831  0.2824  0.2816  0.2809  -1      -1      -1
-1      -1      -1      0.2854  -1      0.2839  0.2831  0.2824  0.2816  0.2809  0.2801  -1
```

Picture of cost layer file.

It is important that the cost layer raster has the same grid size as the species distribution map rasters! **This means that in all files the number of columns and rows as well as the size of cells should be equal**. It is equally important that all those cells which have data of any of the species (or whatever features) used in the analysis, also have to have a cost value. In other words, all cells that have data for any species need a cost value > 0, otherwise undefined program behaviour may occur. Remember also to use **decimal points, not commas**, in all the input files.

Run settings for integrating land cost in the analysis

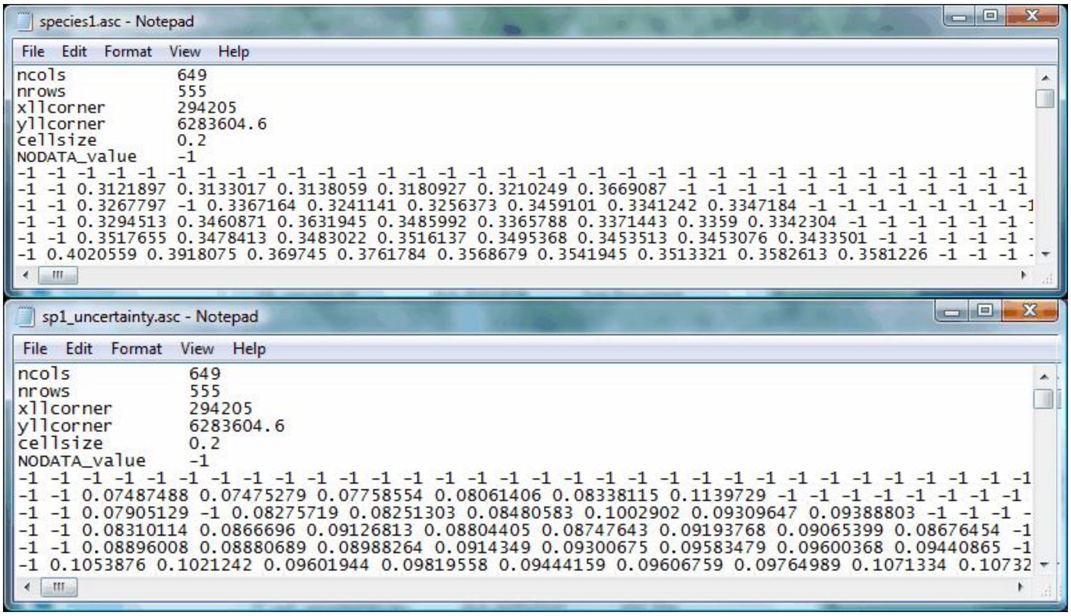
When using the cost layer, remember to type in to your Run settings -file "`use cost = 1`" (cost option selected) and "`cost file = yourcostfile.asc`" (name of your cost layer file).

3.3.3.7 Distributional uncertainty map layer

To account for uncertainty in species distribution data in your conservation planning with Zonation, you need two types of files: **a set of uncertainty map layers**, one for each species, and **an uncertainty weights file**.

Uncertainty map layer is a standard GIS raster file (.ascii or .img) of uncertainties in species occurrence. These files are needed if you are including the uncertainty in species (or whatever feature) distributions (section 2.5.1.) into your analysis, where you need one uncertainty map layer for each species used in the analysis. The file includes all basic raster information as explained in biodiversity feature map files (section 3.3.2.1.) and a matrix of species occurrence uncertainties in each cell (parameter **wsc** in the info-gap uncertainty model). The species-cell-specific uncertainty value wsc can be any measure of error in prediction, or any uncertainty about whether the species will persist there, or a combination of those, as long as the data of all species is in same format. Zonation can treat the values as uniform errors or proportional errors. Uniform error is the default setting and works for most of the data sets, but in some cases it is more appropriate to use proportional errors (see e.g. Ben-Haim 2001).

For example, if confidence intervals are available for the probabilities of occurrence of species A in a given cell, the uncertainty value for species A in that same cell can be the size of the (lower half of the) respective confidence interval. Or the probabilities of anthropogenic threat, e.g the uncertainty of occurrence of species in a given cell due to human activities in the near future, can be used as uncertainty value. Or both. The higher the uncertainty value, the greater the risk that the species does not actually occur there (although the species distribution data might suggest so). Thus, an uncertainty value of 0 indicates that the observed occurrence of a species A in a given cell is trusted to be completely accurate.



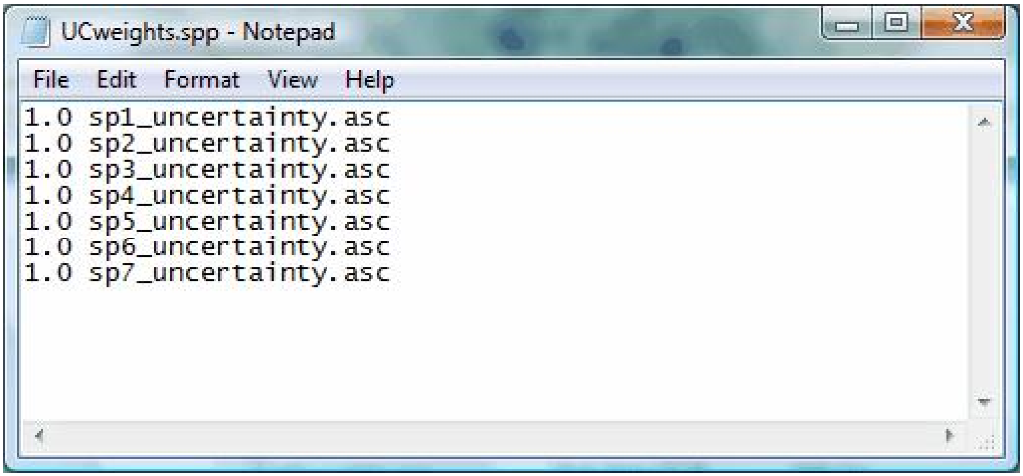
```
ncols         649
nrows         555
xllcorner     294205
yllcorner     6283604.6
cellsize      0.2
NODATA_value  -1
-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1
-1 -1 0.3121897 0.3133017 0.3138059 0.3180927 0.3210249 0.3669087 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1
-1 -1 0.3267797 -1 0.3367164 0.3241141 0.3256373 0.3459101 0.3341242 0.3347184 -1 -1 -1 -1 -1 -1 -1 -1
-1 -1 0.3294513 0.3460871 0.3631945 0.3485992 0.3365788 0.3371443 0.3359 0.3342304 -1 -1 -1 -1 -1 -1 -
-1 -1 0.3517655 0.3478413 0.3483022 0.3516137 0.3495368 0.3453513 0.3453076 0.3433501 -1 -1 -1 -1 -1 -
-1 0.4020559 0.3918075 0.369745 0.3761784 0.3568679 0.3541945 0.3513321 0.3582613 0.3581226 -1 -1 -1 -
```

```
ncols         649
nrows         555
xllcorner     294205
yllcorner     6283604.6
cellsize      0.2
NODATA_value  -1
-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1
-1 -1 0.07487488 0.07475279 0.07758554 0.08061406 0.08338115 0.1139729 -1 -1 -1 -1 -1 -1 -1 -1
-1 -1 0.07905129 -1 0.08275719 0.08251303 0.08480583 0.1002902 0.09309647 0.09388803 -1 -1 -1 -
-1 -1 0.08310114 0.0866696 0.09126813 0.08804405 0.08747643 0.09193768 0.09065399 0.08676454 -1
-1 -1 0.08896008 0.08880689 0.08988264 0.0914349 0.09300675 0.09583479 0.09600368 0.09440865 -1
-1 0.1053876 0.1021242 0.09601944 0.09819558 0.09444159 0.09606759 0.09764989 0.1071334 0.10732
```

Picture showing both the species distribution layer (species1.asc) and the uncertainty layer (sp1_uncertainty.asc) for species 1. During the distribution discounting process the value in each cell of the distribution layer will be discounted by a multiple of the corresponding value in the uncertainty layer.

It is compulsory that the uncertainty layer raster has the same grid size as the species distribution map rasters! **This means that in all files the number of columns and rows as well as the size of cells should be equal**. Also, for each feature, all cells with occurrence data should have a respective uncertainty value. Remember also to use **decimal points, not commas**, in all the input files.

The uncertainty weights file contains a list of names for all distributional uncertainty map layers that correspond to feature layers. This file always contains two columns:



```
1.0 sp1_uncertainty.asc
1.0 sp2_uncertainty.asc
1.0 sp3_uncertainty.asc
1.0 sp4_uncertainty.asc
1.0 sp5_uncertainty.asc
1.0 sp6_uncertainty.asc
1.0 sp7_uncertainty.asc
```

Picture of uncertainty analysis weights file.

1. **Species-specific weights in the uncertainty analysis**. With these you can stress the accuracy of occurrence of a certain species (e.g. very rare species). The higher the weight, the more strongly the program prefers cells with low uncertainty. Species-specific weights can have any positive value larger than 0. If no species-specific weighting of uncertainty is used (as is most commonly the case), this should be set as 1.0 - all equal.

2. Name of the **distributional uncertainty map layer for the feature**. If your uncertainty maps are in different directory than your weights file, remember also to type the correct path in front of the names. Note that the uncertainty layers for species have to be in the same order as the biodiversity feature maps in the biodiversity feature list file - these files are linked to each other solely via the order of listing in the two files.

Note that there shouldn't be any empty rows at the end of the uncertainty analysis weights file. This is because the program might interpret these as empty values or files that just don't have any names. Remember also to use **decimal points, not commas**, in all input files.

Run settings for including uncertainty of distributions in your analysis

To set up an analysis to account for uncertainty in species distributions, you must have the IG-alpha parameter unequal to zero in the fourth last column of your Zonation call (see section 3.3.2.1.). When **uncertainty** about distribution is considered **a negative thing** (see distribution discounting, section 2.5.1.), **the value of α should be positive**. If uncertainty is considered **a positive thing** (see opportunity analysis, section 2.5.2.), **the value of α should be negative**.

Additionally, you need to type in to your Run settings -input file the following rows:

`[Info-gap settings]`

`Info-gap proportional = 0` OR `1` depending on whether uncertainty values are uniform errors (value = 0) or proportional errors (value = 1). If you do not type in this row, the value will be set to 0.

`use info-gap weights = 1` (uncertainty analysis selected)

`Info-gap weights file = yourweightsfile.spp` (the name of your uncertainty weights file)

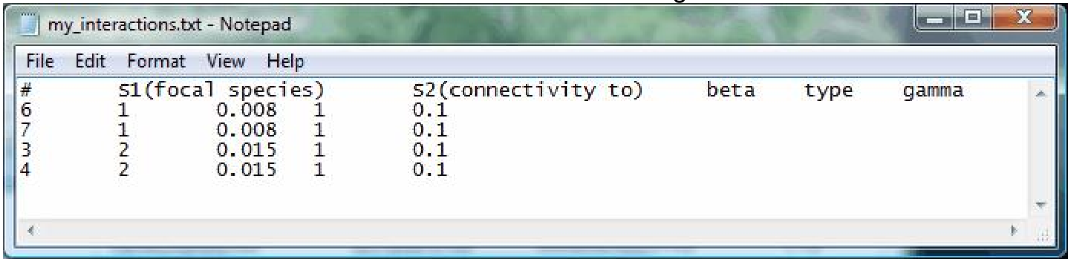
3.3.3.8 Interactions definition file

Species interactions definition file is a text file which defines what interactions between species (or other input features) are to be included in the analysis, implementing methods described by Rayfield et al. (2009). This file is only needed if you want to include ecological interactions into your analysis.

To include species interactions to your analysis you need to first create the interactions definition file. This file tells the program which species (or other input features) are interacting and to what extent. Also, when preparing your biodiversity feature list file, please keep in mind the following points:

- **Every interaction changes the loaded, original distribution layer**. For example, lets assume you have two species (A and B), and you wish to run an analysis with the two original distributions plus the connection of species A distribution to species B distribution. To do this, you need to list species A twice in your species list file: The first layer will be used as it is (original distribution), and the second one will be transformed based on its connectivity to species B. If species A were listed only once, the landscape ranking would be done only based on species B distribution and species A connectivity to species B. Thus species A original distribution would not be included.
- **Make sure that you are not using already transformed layer to transform other layers**. Unless you absolutely want to. However, in that case the interpretation of results is outside the scope of this user manual.
- **Note that every file listed in the species list file will be used for landscape ranking**. If you want to use a layer to transform another layer, but not to be included into the analysis itself, you can do this by setting the weight of the layer to zero.

The structure of the interaction definition file is as in the figure below:



```
#       S1(focal species)       S2(connectivity to)     beta    type    gamma
6       1       0.008   1       0.1
7       1       0.008   1       0.1
3       2       0.015   1       0.1
4       2       0.015   1       0.1
```

The first column (S1) is the focal species, the map of which is transformed. The number in column one is the ordinal number of the species, referring to the order of species in the biodiversity feature list file; in the example above, the first row indicates that species map number 6 is transformed by connectivity to map number 1, using distance dependence (beta, β_k) 0.008. Beta is equivalent to the alpha value that is given in the species list file. Thus, beta is calculated using the same formula

$$\beta = \frac{2 * [\text{Cell size in km}]}{[\text{Distance dependence (km)}] * [\text{Input cell size}]},$$

where the species-specific measure for landscape use refers to the distance, to which focal species S1 in cell *i* can interact with species S2 in the surrounding area of cell *i*.

The fourth column, type, specifies whether this interaction is a positive (e.g. resource-consumer), or a negative one (e.g. competition) (see section 2.6.). The fifth column gives the value of gamma, which in turn defines how the value of connectivity between S1 and S2 changes when moving away from the focal site. For more detailed explanation about gamma, see section 2.6. By default, gamma should be 1.0.

Run settings to include species interactions in your analysis

If you are running the program from command prompt, type into your Run settings -file "`use interactions = 1`" (option selected) and "`interaction file = myfile.txt`" (name of your interaction definition file).

3.3.3.9 Removal mask layer

Zonation v.3.0 includes an improved mask file function. **Please note that a removal mask layer composed to this new format is no more compatible with Zonation v.2.0.** Please note that there are two different mask types for different purposes. Do not confuse them.

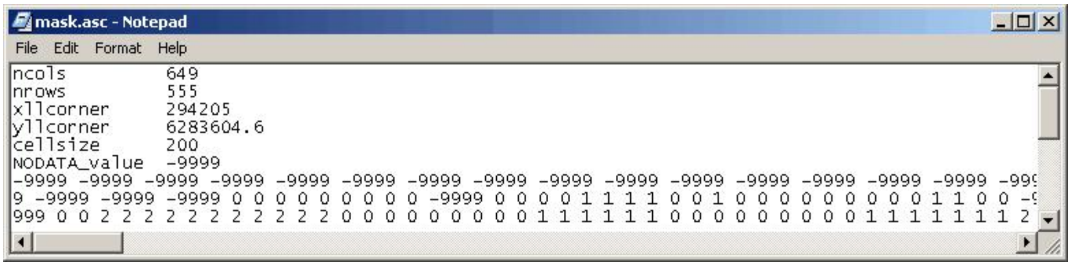
The removal mask layer is a GIS raster file which determines the removal hierarchy of the edge cells. The typical uses of the mask layer include replacement cost analysis (section 2.6), and conservation prioritization when some predetermined information about land zoning exists. This feature is different and improved from its precedent in Zonation v.2.0. The earlier mask file feature had problems with

- slightly awkward coding for areas masked in or out
- only having three levels: masked out, normal and masked in
- cells causing zero marginal loss could end up treated wrong due to the way the feature was implemented

The new mask file in Zonation v.3.0. works as follows:

- It is a raster file (.asc or .img or .tif) where integer values are assigned to cells.
- The cell-specific integer is the "mask level" for the cell.
- Cells with lower mask levels (small numbers) are removed before cells with higher mask levels (large numbers), thus making possible a forced multi-level hierarchy in the analysis (see Lehtomäki et al. 2009).
- There can be an arbitrary number of levels, 0, 1, 2, 5, 7, 100, 1500, 100000, whatever integers; the levels need not be consecutive.

Cells with low mask levels are removed first, and they may be, for example, undesirable for conservation (e.g. built-up areas, private areas, areas ear-marked for residential building or commercial fishing etc.) or they may have any other reason to be primarily excluded from the final solution. The cells with high mask levels that are removed last may, for example, have a special conservation value or they may already be ear-marked for conservation. These cells will be removed only after there are no more cells with lower mask level values left, and thus become included into the top fraction of the solution.



```
ncols         649
nrows         555
xllcorner     294205
yllcorner     6283604.6
cellsize      200
NODATA_value  -9999
-9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -99
9 -9999 -9999 -9999 0 0 0 0 0 0 0 0 0 -9999 0 0 0 0 1 1 1 1 0 0 1 0 0 0 0 0 0 0 0 0 1 1 0 0 -
999 0 0 2 2 2 2 2 2 2 2 2 2 2 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 2
```

Picture of removal mask layer file.

It is compulsory that the mask layer raster has the same grid size as the species distribution map rasters! This means that in all files the number of columns and rows as well as the size of cells should be equal. It is equally important that all those cells which have data of any of the species used in the analysis (that is to say, the cells that are NOT marked as "no data" in all species distribution files), also have a mask value.

Run settings for using a removal mask layer

When using the removal mask layer, remember to type in to your Run settings -file, under the `[Settings]` heading:
`run mask = 1` (mask option selected)
`mask file = yourmaskfile.asc` (name of your removal mask layer file).

Note that use of removal mask layer is likely to result a biologically non-optimal solution as the program is not allowed to remove cells only based on their conservation value. See **replacement cost analysis** (sections 2.7. and 5.2.3.) for the analysis of the suboptimality of masked solutions; with areas forcibly masked to low or high ranks of the solution.

3.3.3.10 Analysis area mask

This raster file (.ascii or .img or .tif) indicates areas to be used in the analysis. If analysis area mask is used, Zonation will only take into account those cells that are indicated by this file. Please note that this mask is not suitable for forcing areas in to the top or bottom fraction of the solution! If that is what you want to do, your choice is removal mask layer (section 3.3.3.9.).

The area mask file will be applied to all input raster files. It can be used for
(i) Forcing alignment of data
(ii) Cutting out areas that are not needed in analysis, e.g. areas outside the country, and thereby saving memory and allowing more features to be analysed.
(iii) Targeting analysis to a subsection of the landscape without needing to develop a full new set of files - just the area mask file will suffice.

This file includes all basic raster information as explained in species distribution map files, followed by a matrix, where cells are categorized as follows: Cells with value >0 are included in the analysis, whereas cells with zeroes or "missing data" values, e.g. -1 are excluded. A simple operational area mask file will have "1"s for cells that should be analysed and "0"s for cells that are outside the area of interest.

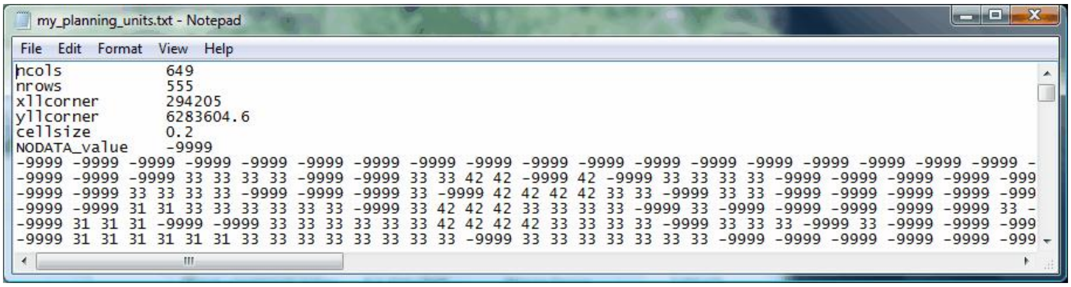
It is compulsory that the analysis area mask raster has the same grid size as the species distribution map rasters! **This means that in all files the number of columns and rows as well as the size of cells should be equal**.

Run settings for using analysis area mask

To use analysis area mask in your analysis, type in to your Run settings -file "`mask missing areas = 1`" (analysis area mask is used) and "`area mask file = your_areamask.asc`" (name of your area mask raster).

3.3.3.11 Planning unit layer

The planning unit layer file is a standard GIS raster file (must be .asc ASCII file) containing integer numbers. This file includes all basic raster information as explained in biodiversity feature map files (see section 3.3.2.1.) and a matrix where the number given for a cell identifies the planning unit that the cell belongs to. Planning unit numbers must be positive integers, but they need not be sequential, they could be e.g. 1, 2, 5, 12, 1010, … Planning units may be used to model the situation when, e.g., land ownership dictates that certain groups of grid cells should be treated as distinct units. Or, with directed connectivity, the planning units could corresponds to hydrologically linked catchment areas. When the planning unit layer is in use, the entire planning unit is removed simultaneously. The cell removal rules operate as before, but they operate on value aggregated across the planning unit. Also, the cost of the planning unit is taken as either the summed cost of cells (if the cost layer is used) or as the area of (number of cells in) the planning unit if costs are not used. Note that each planning unit does not need to be spatially continuous; a planning unit may consist of a scattered collection of cells.



```
ncols         649
nrows         555
xllcorner     294205
yllcorner     6283604.6
cellsize      0.2
NODATA_value  -9999
-9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -9999 -
-9999 -9999 -9999 33 33 33 33 -9999 -9999 33 33 42 42 -9999 42 -9999 33 33 33 33 -9999 -9999 -9999 -9999 -999
-9999 -9999 33 33 33 33 -9999 -9999 -9999 33 -9999 42 42 42 42 33 33 -9999 33 33 -9999 -9999 -9999 -9999 -999
-9999 -9999 31 31 33 33 33 33 33 33 -9999 33 42 42 42 33 33 33 33 -9999 33 -9999 -9999 -9999 -9999 -9999 33 -
-9999 31 31 31 -9999 -9999 33 33 33 33 33 33 42 42 42 42 33 33 33 33 -9999 33 33 33 -9999 33 -9999 -9999 -999
-9999 31 31 31 31 31 31 33 33 33 33 33 33 33 33 -9999 33 33 33 33 33 33 33 -9999 -9999 -9999 -9999 -9999 -999
```

A planning unit number should be defined for every cell that has species data, if not, some kind of error condition is likely to occur. It is not harmful to have planning units extending outside the area with species data, the critical bit is that all locations with species data are covered with planning units.

Use of large planning units will automatically cause a decrease in the quality of results. The reason for this is that large planning units will probably contain areas that are both good and bad for conservation. Consequently, the performance curves will suggest lower protection levels than what can be obtained if selection is based on individual grid cells. With respect to the computation time needed by Zonation, use of planning units is likely to cut computation times. The reduction will be the larger the more cells are grouped into planning units.

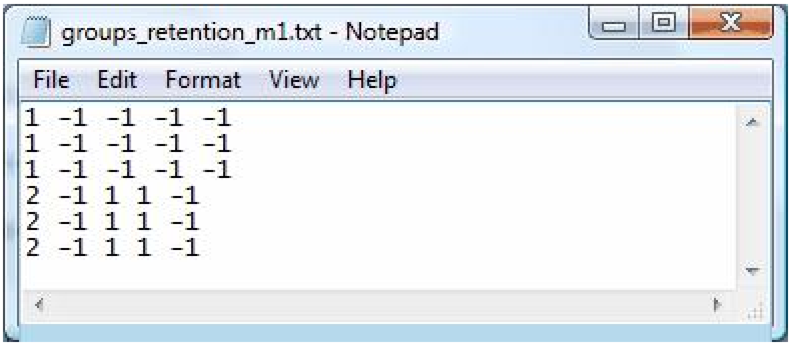
Run settings for planning units

When using planning units, remember to type into your Run settings -file "`use planning unit layer = 1`" (planning unit option selected) and "`planning unit layer file = my_plu_layer.asc`" (name of your planning unit layer file). Note that when using planning units as input data, the program automatically sets warp factor to one.

3.3.3.12 Groups file

The groups file allows linking input features to groups, upon which various operations are done. The groups file has as many rows as there are input features. Columns of the file define specific groupings, one per column. All numbers in this file are integers; they are ordinal numbers that group features. Note that giving a -1 for any group for a feature indicates that the feature is not grouped with respect to that particular criterion. After the descriptions of file columns, there is an example that will hopefully illuminate the operation of the groups file.

Assignments to different types of groups for each feature are listed in the groups file. There is one row for each feature. Reading the file, Zonation refers to features in the species list file, so the order of the features in the groups file must match that in the species list file. For each feature, there are five columns, one column for each group type:
column 1: output group
column 2: condition group
column 3: retention groups
column 4: retention mode
column 5: local edge correction group



```
1 -1 -1 -1 -1
1 -1 -1 -1 -1
1 -1 -1 -1 -1
2 -1 1 1 -1
2 -1 1 1 -1
2 -1 1 1 -1
```

An example of a groups file.

Presently, the meanings of the five columns are:

i) Column 1, **output group** can be useful if you are using mixed sets of biodiversity features and want to assess representation for each set separately. Zonation will output mean, minimum and maximum representation curves for each of the groups in a separate groups output file (section 3.4.3.). You can, for example, assign output groups to separate a) different higher taxa: birds, mammals, plants etc., b) community and species features, c) negatively and positively weighted features, d) habitat quality and connectivity layers, or creative combinations of any of the aforementioned. The information could, of course, be extracted manually from the standard representation curves file (section 3.4.1.). Using output groups can save you some manual work and make interpretation of the results more straightforward.

ii) Column 2, **condition group** defines linkage of features to condition groups.

There is another file (the condition layers list file; section 3.3.3.14.), which specifies which condition number links to which physical file. If the condition group for a feature is, say, 3, then the feature is linked to the third layer in the condition layers list file.

The condition column is different from the output group column in that only the one column of information is needed to define the output group. The condition feature additionally needs the condition layers list file and the condition layers themselves.

iii) Column 3, **retention group** defines linkage of features to retention groups.

The retention group is completely analogous to the condition group. The linkages to condition and retention features are structurally identical.

There is another file (the retention layers list file; section 3.3.3.15.), which specifies which retention number links to which physical file. If the retention group for a feature is, say, 3, then the feature is linked to the third layer in the retention layers list file.

iv) **Retention mode** can be 1 = stop loss or 2 = management intervention (see section 2.10. for descriptions).

v) Column 5 refers to an analysis feature that is not available in the current version. This column should always be assigned a value of -1 for all biodiversity features listed.

Example

If, for example, the group file has a row

```
3  1  -1  1  -1,
```

it means that the feature belongs to output group number 3, and condition group 1. There is no retention group specified (-1 in column 3), meaning that the retention mode of column 4 is a dummy value. Column 5 is not used either.

Going back to condition, this means that in the condition file (section 1.3) there must be a row

```
1  cond_group_1_grid_file_name.asc.
```

This, in turn, means that condition group 1 is linked to raster grid file cond_group_1_grid_file_name.asc.

Examples of using output groups are provided in Exercises 10 and 11.

Run settings for using groups file in your analysis

If you are running the program from command prompt, type into your Run settings -file " `use groups = 1`" (option selected) and "`groups file = mygroupsfile.txt`" (name of your groups file).

3.3.3.13 Alternative land uses layer

In Zonation v.3.0, it is possible to consider multiple opportunity costs of conservation by including layers that describe landscape suitability for competing land uses. The alternative land uses layers are compiled and treated precisely as the normal biodiversity feature layers (section 3.3.2.1.) and listed in the biodiversity feature list file (section 3.3.2.2.). The only difference is that the alternative land use layers are given negative weights instead of positive ones in the first column of the list file. If you are using target-based planning as your cell removal rule, the value you enter to the fifth column of your biodiversity feature list file denotes the fraction of the competing land use that you want to have excluded from your solution. Several alternative land uses can be entered in a single analysis, each of them as a layer of their own.

It is compulsory that all feature grids, including negatively weighted ones, have the same numbers of columns and rows.

3.3.3.14 Condition layer

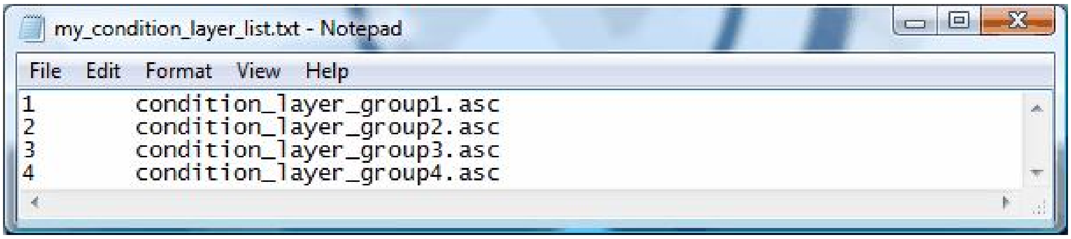
Condition layers are needed when landscape condition is included in the analysis to account for past habitat loss or degradation. They are raster grid layers (.ascii or .img) that describe the fraction of suitable habitat or occurrences that remains for a group of species or other biodiversity features in each grid cell, relative to some historical baseline. The condition values can vary between 0.0 and 1.0. A value of 1.0 indicates pristine condition, in which the habitat suitability or species occurrence has not degraded. A value of 0.0 indicates a completely degraded condition. Any negative values (missing data) will be treated as zeroes.

It is compulsory that the condition layers have the same grid size as the biodiversity feature map rasters!

When a condition grid layer is included in a Zonation run, the value of a cell for a species (or other biodiversity feature) in the distribution map layer will be multiplied by the condition value of the cell for the condition group in which the feature belongs to. The condition grid layers need to be accompanied by two files: a groups file and a condition layer list file.

Groups file (section 3.3.3.12.) assigns the biodiversity features to condition groups based on, for example, their habitat preferences. Condition groups are assigned in column 2 of this file.

Condition layer list file links condition groups to the condition grid layers. This file has two columns:
Column 1 gives the number of the condition group. These numbers refer to those assigned to biodiversity features in the groups file.
Column 2 has the name of the condition grid layer for that group of features.


my_condition_layer_list.txt - Notepad
File Edit Format View Help

```
1       condition_layer_group1.asc
2       condition_layer_group2.asc
3       condition_layer_group3.asc
4       condition_layer_group4.asc
```

An example of a condition layer list file.

Run settings for landscape condition analysis

To include landscape condition in your analysis, make sure you have the following rows typed in your run settings file, section [settings]:

```
use groups = 1
groups file = mygroupsfile.txt (name of your groups file)
use condition layer = 1
condition file = condition_layers_list.txt
```

3.3.3.15 Retention layer

A retention layer is a raster grid (.ascii or .img, etc.) that describes the fraction of local occurrences or habitat suitability that would be retained for a group of species (or other biodiversity features) even in the absence of conservation action. Values in the cells can vary. A value of 0 indicates that in the absence of conservation, the cell will lose all of its biodiversity value. A value of 1 indicates that the cell will be retained as it is and no loss would occur. A value of >1.0 indicates that condition in the cell will improve with management; values >1.0 are only appropriate for the 'management gain' mode of analysis, mode 2.

It is compulsory that retention layers have same dimensions as the biodiversity feature map rasters!

When a retention layer is included in a Zonation analysis, it transforms the occurrence of the feature to expected loss by multiplying the value of the feature in the cell of the distribution grid by the (1-retention level) of the same cell for the retention group the species belongs to.

The retention layers need three files to be operational:

A groups file (section 3.3.3.12.) assigns the biodiversity features to retention groups. The groups file also defines the mode of retention, which can be either that (1) reduced loss of habitat occurs or that (2) the quantity or quality of habitat is actually improved through management intervention. Retention groups are assigned in column 3 and retention modes in column 4 of this file.

A retention layer list file links retention groups to the retention grid layers. This file has two columns:

Column 1 gives the number of the retention group. These numbers refer to those assigned to biodiversity features in the groups file.
Column 2 has the name of the retention grid layer for that group of features.

The retention layers themselves. The retention layer describes retention in terms of the fraction of cell condition retained even in the absence of conservation. A value of 1.0 indicates no change in condition even in the absence of conservation. A value of 0.0 (mode 1) would indicate total loss in the absence of conservation. A value of, say, 1.2 (mode 2), would imply a 20% management gain assuming the site is targeted for conservation.

An example of a retention layer list file.

Often, the most meaningful way to include landscape retention in the analysis is to use it in combination with species representation in the protected areas only. To do this, you need to duplicate your biodiversity features in the biodiversity feature list file (see section 5.3.5. for the full setup). The first copy of the layers will be used to model representation and the second one will model loss in occurrence if the cell is not selected to the protected area network. The relative weights for representation and retention are defined by parameter β which is given in the run settings file (see below). The higher the value, the more emphasis is given to retention in the analysis.

When landscape condition and retention are both applied in the same analysis, the occurrence values of features in cells are first transformed by condition and then by retention levels.

Run settings for landscape retention analysis

To include landscape retention in your analysis, adjust your run settings file to include the following lines:

`use groups = 1`
`groups file = mygroupsfile.txt` (name of your groups file)
`use retention layer = 1`
`retention file = my_retention_layer_list.txt`
`retention layer relative weights =` value for β, a decimal multiplier for the retention layer weights for balancing between representation and retention

3.3.3.16 Administrative units analysis files

To run an analysis where you consider conservation priorities over multiple administrative regions, you need three additional files:
(i) An administrative units description file that defines local and global factors for multiplying weights of species or other biodiversity features,
(ii) an administrative units map raster that assigns cells of the whole planning region to administrative subregions and
(iii) administrative units weights matrix defining weights for each biodiversity feature in each administrative region.

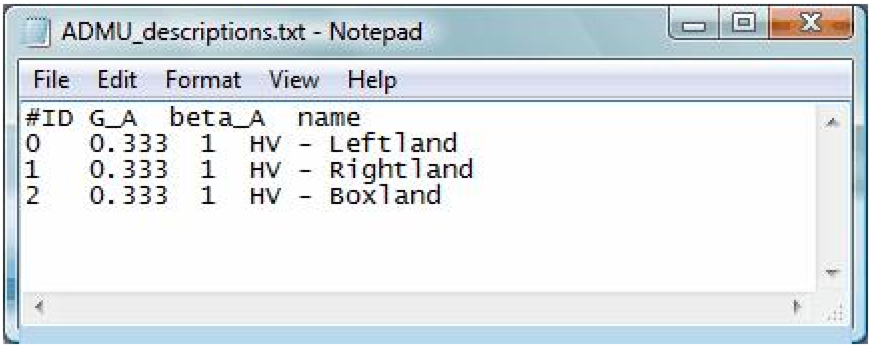
(i) **Administrative units description file** is a list of administrative regions and their weights. The first row is a header row for column names. The file has four columns:

Column 1 has the identification number of the administrative region (these should be positive integers). The values are linked to the administrative units map raster and should match those.

Column 2 has the region-specific global weight G_A. This number is the global priority given for this subregion, Effectively, conservation value aggregated from this region becomes multiplied by this number; a high number elevates the priority given to the region relative to other regions. Note that in the strong variant (ADMU mode 2) large global weights for the areas increase the relative importance of the local component because conservation values of the regions are summed (loss of conservation value is computed from 1+N components, where 1 is for global analysis and N is for N subregions). The weights in this column can be standardized to sum to one to keep the balance between global and local considerations directly dependant on parameter p.

Column 3 has the region-specific local weight balance between local and global feature weights. This number is bounded between 0 and 1. It gives the balance of local vs. global considerations (feature weights) in local decision making. A number of 1 here means that the region only cares of its own local priorities (w^L_{jA}). A zero here means that the region is willing to go with global priorities (w^G_j), thus foregoing local opinions of what is important.

Column 4 gives the name or ID of the region for output purposes.



```
#ID G_A   beta_A  name
0   0.333  1   HV - Leftland
1   0.333  1   HV - Rightland
2   0.333  1   HV - Boxland
```

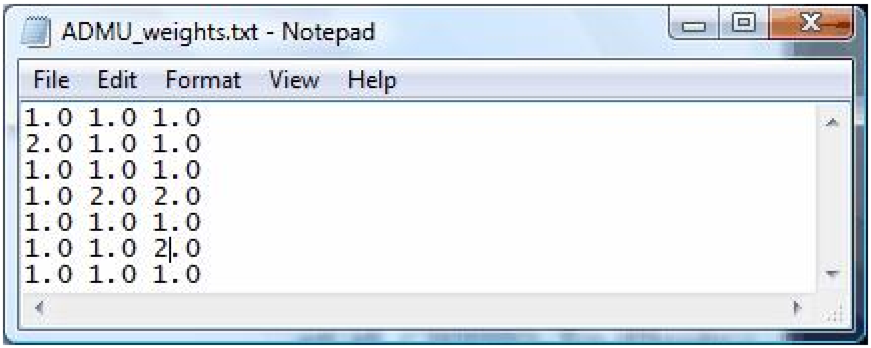
An example of an administrative units description file.

(ii) **Administrative units map raster** is an .asc ascii raster grid that contains integer numbers. It defines the division of grid cells to the administrative regions. Each administrative region has its own identification number, defined in the administrative units description file. Cells in this map raster should contain the identification numbers of the administrative regions they belong to.

It is possible to use the administrative units in combination with planning units. If you do this, please make sure that the planning units should not go across the borders of administrative regions.

It is compulsory that the ADMU raster has the same dimensions as the feature grids. The ADMU number should be defined for all locations for which feature data is entered. ADMU number can be defined for locations without any feature data, i.e., one can use an ADMU layer based on simple polygons even if the feature data only covers parts of them.

(iii) **Administrative units weights matrix file** defines the a priori local weights of species or other biodiversity features used in the analysis. Here, biodiversity features are in rows and administrative regions in columns. The local weights could reflect the regional conservation priorities or policies. The order of biodiversity features in the matrix file should match that of the biodiversity feature list file. This file has no header row.



```
1.0 1.0 1.0
2.0 1.0 1.0
1.0 1.0 1.0
1.0 2.0 2.0
1.0 1.0 1.0
1.0 1.0 2.0
1.0 1.0 1.0
```

An example of an administrative units weights matrix.

Run settings for conservation prioritization over multiple administrative regions

For an administrative units analysis, add the following lines to your run settings file (using your own file names of course):

```
[Administrative units]
use ADMUs = 1
ADMU descriptions file = my_ADMU_descriptions.txt
ADMU layer file = ADMUs_distribution_map.asc
ADMU weight matrix = ADMU_weights_matrix.txt
```

`calculate local weights from condition = 1` Presently a non-operational dummy parameter.

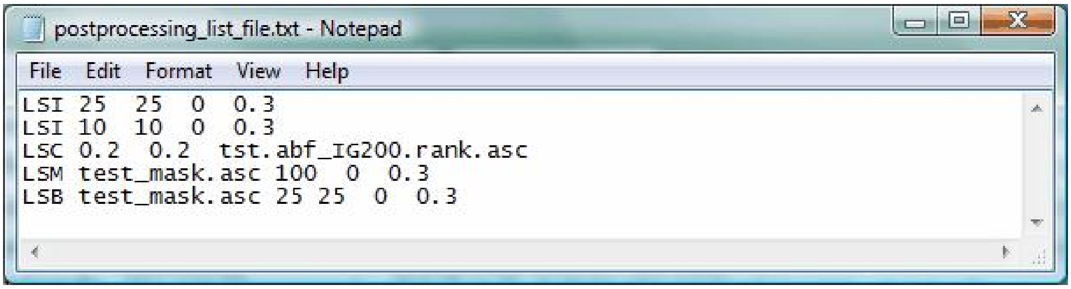
`ADMU mode = 1` OR `2`, depending on whether you want to enforce local representation for all features (mode 2) or not (mode 1; see section 2.12.)

`Mode 2 global weight = 0.5` A parameter specifying the balancing of global and local conservation value when using ADMU mode = 2. This parameter is bounded between zero (only local considerations) and one (only global considerations influence value).

See section 5.3.8. for more detailed instructions for the analysis setup.

3.3.3.17 Automated post-processing file

Automated post-processing file describes post-processing analyses that are called automatically after the main computations have finished. The file contains a list of analyses to be done along with their parameter settings. Below is an example of what post-processing calls look like (there is one row per call in a text file; there can be multiple calls of each type and the calls need not be in any particular order):


postprocessing_list_file.txt - Notepad
File Edit Format View Help

```
LSI 25  25  0  0.3
LSI 10  10  0  0.3
LSC 0.2  0.2  tst.abf_IG200.rank.asc
LSM test_mask.asc 100  0  0.3
LSB test_mask.asc 25 25  0  0.3
```

Run settings for automated post-processing

To enable automated post-processing analyses in batch operation, type into your Run settings -file, section [Settings], "`post-processing list file = ppa_list_file_name.txt`" (name of your post-processing file).

Four different post-processing analyses can be computed:

i) **Landscape identification analysis, identified by analysis type LSI**
This option allows identification of separate management landscapes based on the distance and similarity in species composition between two sites. Spatially distinct areas (consisting of multiple grid cells) in a Zonation solution can be classified into management landscapes. An area is joined to a landscape if it is spatially close enough and similar enough with respect to species composition to *any* other distinct area in the same landscape. Landscape identification is done for a specific top fraction of a Zonation ranking.

To perform a landscape identification analysis, you need to define four parameters:

1. First define the **percentage of landscape**. This will determine how large part of the entire landscape will be included in the classification, including areas from the top fraction of the landscape. Note: this number is given as percents. E.g. value of 20 includes the best 20% area from the landscape to the solution.

2. Define **inclusion minimum**, which in turn determines how highly ranked cells must be included in each management landscape. E.g. value of 10 means that each management landscape has to contain at least one cell which belongs to the top 10% fraction of the whole landscape. Note that if the inclusion minimum is equal to (or larger than) the **percentage of landscape**, all spatially separate areas will be joined into the management landscapes, where as if the inclusion minimum is smaller, then management landscapes only with sufficiently high ranked core areas are included.

3. Give the nearest neighbor **maximum distance** (in cells!), which is allowed between spatially discrete patches that are included in the same management landscape. E.g. a maximum distance of 0 would mean that all separate groups of cells are identified as unique management landscapes.

4. Give the **maximum difference in species composition.** This determines how much the species compositions between two cells are allowed to differ (in terms of relative densities) for them to be joined to the same landscape. A value of 0

indicates that the species composition in two patches is identical. Value of 1 indicates that the difference in relative density between two patches is on average log10 across species. E.g., a maximum difference of 0.2 means that on average two species out of ten have a 1-log difference in their density or that one tenth of all the species have a 2-log difference. For more details, see Moilanen *et al.* (2005).

Thus, if you want to identify management landscapes within your solution, you need the following line in your automated post-processing file:

```
LSI  percent1  percent2  distance  similarity
```

in which "LSI" tells Zonation to perform the analysis, "percent1" is the fraction of landscape to be included, "percent2" is the inclusion minimum, "distance" is the maximum distance and "similarity" the maximum difference in species composition.

Each landscape identification analysis will produce two output files with extensions .nwout.#.ras.asc and nwout.#.spp_data.txt, described at section 5.2.1. The number # in the file name is generated by the ordinal number of the call to an LSI analysis; the first call produces .nwout.1.ras.asc, the second .nwout.2.ras.asc and so on. The beginning of the LSI output files will be the general output file name you have given in the Zonation call.

ii) **Landscape comparison LSC**

Solution comparison calculates how much two solutions overlap with each other and what is the average difference in the cell removal order. The comparison is always made between the present solution and an older solution by using the rank.asc -files of both solutions as input files. For landscape comparison, you need the following line in your automated post-processing file:

```
LSC  fraction_of_present_solution
fraction_of_comparison_solution  comparison.rank.asc  output.asc
```

Here, "LSC" tells Zonation to perform the analysis. "fraction_of_present_solution" defines how large (top) fraction of the present solution is accounted for in the comparison. The value can vary be between 0 an 1. "fraction_of_comparison_solution" defines how large (top) fraction of a previous solution is accounted for in the comparison. Again, the value can vary be between 0 an 1. "comparison.rank.asc" points to the rank output file of the previously computed solution that you want to compare your new solution with. "output.asc" denotes the name of the output raster (section 3.5.1.3.) that shows the overlapping areas.

iii) **Landscape identification for masked subregion of landscape, LSM**

This analysis is, effectively, an LSI analysis done to an externally specified subregion of the landscape.

With this command, you can identify management landscapes within subregions of the full landscape. This subregion, could, for example, be areas owned by a particular land owner. The mask file identifies the areas you are interested in - do not confuse this mask with the mask used in landscape priority ranking - these would typically be two different mask files.

To identify management landscapes for masked areas, add the following line to your post-processing file:

```
LSM   mask_file  percent  distance similarity
```

in which "LSM" calls the analysis and "mask_file" is the name of your raster showing areas of interest (prepare this file so that areas of interest have number >=1 and the rest of the grid cells are missing data, e.g., -1). "Percent" gives the inclusion minimum (see LSI), and "distance" and "similarity" are as in LSI.

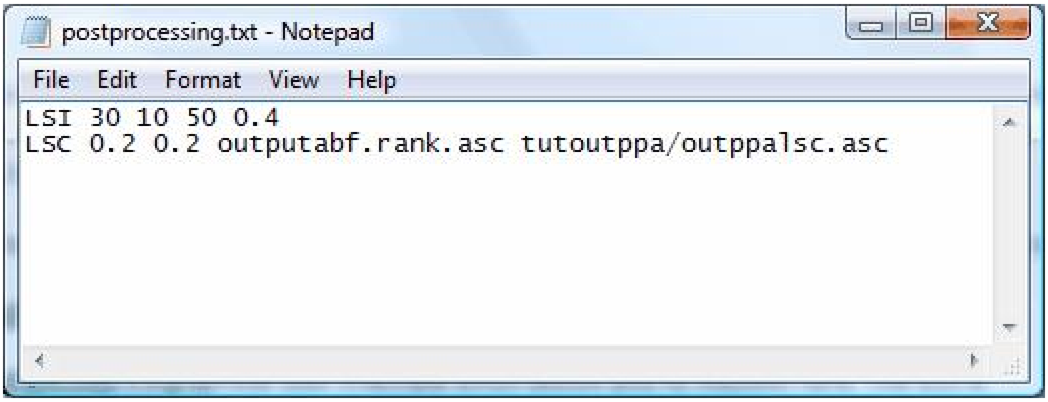
The output of the LSM analysis is like that of the LSI analysis, with each new call adding one to the output file name.

iv) **Landscape identification for top fraction inside masked areas LSB**

With LSB, you can combine a top fraction analysis and a mask file to choose areas used in identification of management landscapes. It is different from LSM in that LSM uses all areas indicated by the mask , but LSB only uses a given top fraction within masked areas. The line to type for LSB is

```
LSB   mask_file  fraction1 fraction2  distance  similarity
```

Here, "mask_file" is as for LSM analysis, and the rest of the parameters are as for LSI.



```
LSI 30 10 50 0.4
LSC 0.2 0.2 outputabf.rank.asc tutoutppa/outppalsc.asc
```

An example of a post-processing file.

3.4 Standard Zonation output

Next we will describe the basic output produced by the program. Running Zonation automatically produces two sets of outputs:

1. **Visual output** in the graphical user interface.
2. **File output**. These files will be saved in the same directory as the program unless you have specified another path for your output.

3.4.1 Automated file output

In addition to the visual output, the program automatically produces a number of different output files for each run. Here we describe the ones that are always produced, regardless of analysis variant. Analysis-specific files are described in the next section. In the command line, you have specified the output filename, e.g. "output", which will be used for each of the output files with a varying suffix: output.jpg, output.emf, output.curves.txt, output.prop.asc, output.rank.asc, output.wrscr.asc and output.run_info.

.jpg file

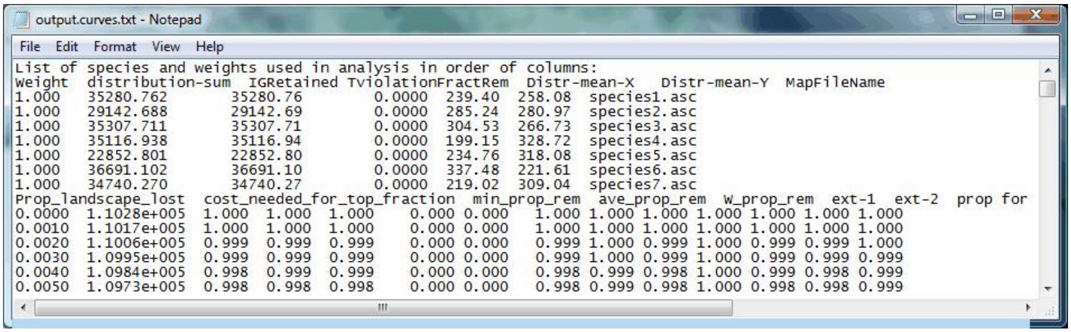
An image of the map of the area illustrating the Zonation results, ranked by using different colors to indicate the biological value of the site. Here the best areas are displayed in red and the worst areas in black, with the "no data" -areas marked as white. See section 4.3.1. for more detailed interpretation of the colors used in landscape ranking.

.emf file

This is an identical image showing your output map, but it has a higher quality compared to the .jpg file. Thus, if you are using any of the images in publications, it is recommended to use this file type. The images can of course be redrawn in GIS from imported .asc raster map files.

.curves.txt file

A text file containing a list of species and the **relative weights** `(Weight)` used in the analysis, together with the initial **sum of species distributions** `(distribution_sum)` and the level of cell removal at which point targets for particular species have been violated. The initial sum of distribution is simply the sum of each species local occurrence levels through the landscape. For example, if your species data is in probabilities of occurrence, this is the sum of probabilities in all cells before any landscape has been removed. The third column shows distribution sum after distribution discounting in an uncertainty analysis `(IGRetained)`. If you are using target-based planning as your cell removal rule, each species has a defined target (e.g. 25% of original distribution) which the program seeks to retain during the removal of landscape. However, as all cells will be eventually removed these targets will be inevitably violated. Thus, in .curves.txt file the program simply reports **when the targets of particular species have been violated**, ie. what fraction of landscape was still remaining when the proportion of species original distribution dropped below the given target `(TViolationFractRem)`. **If target-based planning is not used as a cell removal rule, this column has only dummy values.** After the list you can find columns representing more detailed information of how large proportion of each species' distribution is remaining when landscape is iteratively removed. The last columns output **mean X and Y coordinates** of the original distribution of the feature `(Distr-mean-X and Distr-mean-Y)` and name of the **biodiversity feature map file** `(MapFileName)`.



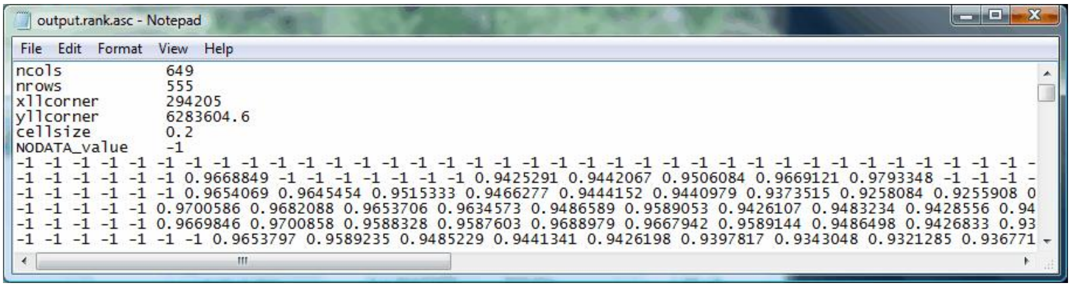
```
List of species and weights used in analysis in order of columns:
Weight  distribution-sum  IGRetained TviolationFractRem  Distr-mean-X   Distr-mean-Y  MapFileName
1.000    35280.762        35280.76         0.0000  239.40  258.08  species1.asc
1.000    29142.688        29142.69         0.0000  285.24  280.97  species2.asc
1.000    35307.711        35307.71         0.0000  304.53  266.73  species3.asc
1.000    35116.938        35116.94         0.0000  199.15  328.72  species4.asc
1.000    22852.801        22852.80         0.0000  234.76  318.08  species5.asc
1.000    36691.102        36691.10         0.0000  337.48  221.61  species6.asc
1.000    34740.270        34740.27         0.0000  219.02  309.04  species7.asc
Prop_landscape_lost  cost_needed_for_top_fraction  min_prop_rem  ave_prop_rem  W_prop_rem  ext-1  ext-2  prop for
0.0000  1.1028e+005  1.000  1.000  1.000    0.000 0.000   1.000 1.000 1.000 1.000 1.000 1.000 1.000
0.0010  1.1017e+005  1.000  1.000  1.000    0.000 0.000   1.000 1.000 1.000 1.000 1.000 1.000 1.000
0.0020  1.1006e+005  0.999  0.999  0.999    0.000 0.000   0.999 1.000 0.999 1.000 0.999 0.999 1.000
0.0030  1.0995e+005  0.999  0.999  0.999    0.000 0.000   0.999 1.000 0.999 1.000 0.999 0.999 0.999
0.0040  1.0984e+005  0.998  0.999  0.999    0.000 0.000   0.998 0.999 0.998 1.000 0.999 0.998 0.999
0.0050  1.0973e+005  0.998  0.998  0.998    0.000 0.000   0.998 0.999 0.998 1.000 0.998 0.998 0.999
```

Picture of output.curves.txt file.

The first column `(Prop_landscape_lost)` gives the proportion of the landscape removed. If you have initially removed some parts of the landscape before running the program (initial removal), the file contains only those areas that are included in the analysis. **The second column** `(cost_needed_for_top_fraction)` shows the cost of remaining landscape. If land costs are not included in the analysis, this column represents the number of cells that is remaining in the landscape. **The third column** `(min_prop_rem)` shows the minimum proportion of species distribution that is remaining in the landscape, thus the situation of the worst-off species. **The fourth column** `(ave_prop_rem)` represents the average proportion over all species. **The fifth column** `(w_prop_rem)` gives the weighted average proportion of all species. **The sixth column** `(ext-1)` shows the average extinction risk of species as landscape is iteratively removed (as calculated from the species-area relation using parameter *z*) and **the seventh column** `(ext-2)` is the weighted extinction risk where species-area extinction risk has been weighted by the species weights. **The following columns** show the proportion of distribution remaining for each species in the same order as the species are listed in the beginning of the file. Note, that for the output file to be readable, the program does not print every step of cell removal (this file only includes a maximum of 10 000 rows).

.rank.asc file

A raster file representing the ranking of the landscape, or in other words, the order of cell removal. The file includes all basic raster information as explained in species distribution map files and a matrix of cell removal order. Here the cells receive a value between 0 and 1. Low values close to zero indicate that the cell has been removed in a early state of the process whereas cells with high value are removed last.



```
ncols         649
nrows         555
xllcorner     294205
yllcorner     6283604.6
cellsize      0.2
NODATA_value  -1
-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -
-1 -1 -1 -1 -1 -1 0.9668849 -1 -1 -1 -1 -1 -1 -1 0.9425291 0.9442067 0.9506084 0.9669121 0.9793348 -1 -1 -1 -
-1 -1 -1 -1 -1 -1 0.9654069 0.9645454 0.9515333 0.9466277 0.9444152 0.9440979 0.9373515 0.9258084 0.9255908 0
-1 -1 -1 -1 -1 0.9700586 0.9682088 0.9653706 0.9634573 0.9486589 0.9589053 0.9426107 0.9483234 0.9428556 0.94
-1 -1 -1 -1 -1 0.9669846 0.9700858 0.9588328 0.9587603 0.9688979 0.9667942 0.9589144 0.9486498 0.9426833 0.93
-1 -1 -1 -1 -1 -1 -1 0.9653797 0.9589235 0.9485229 0.9441341 0.9426198 0.9397817 0.9343048 0.9321285 0.936771
```

Picture of output.rank.asc file.

.prop.asc file

A raster file similar to .rank.asc -file. Here, however, the matrix shows the proportional loss of distribution for that species, that has lost most of its distribution during the landscape removal process. E.g. if a cell has a value of 0.7, it means that after removing that cell all species have at least 30 % of their distribution left (the value 0.7 indicates that one of the species, which is doing worst after removing that particular cell, has lost exactly 70 % of its distribution).

.wrscr.asc file

In addition to the .rank.asc and .prop.asc, a third map is output automatically. This is the .wrscr.asc file, where *wrscr* stands for Weighted Range Size Corrected Richness. This map reports for each cell *i* the quantity

$$wrscr_i = \sum_j w_j q_{ij},$$

where w_j is the weight of species *j* and q_{ij} is the fraction of the distribution of the species in the cell. The measure is simply a sum over species of the weighted fraction of species distributions occurring in the cell, as measured from original input distributions. To illustrate, the cell could have many occurrences of widespread low-weight species. In this case, despite high richness *per se*, the *wrscr*-value would be low compared to another cell which does not contain many species but does have a significant fraction of the entire range of an (endemic) species or two, with relatively high weight.

This map can be used as a scoring value for the cell, which can be useful for example when comparing two cells with a replacement cost value of zero – the cell with the higher *wrscr*-value would be more important. *Wrscr*-values could be used, for example, to inform agro-urban land use planning of the potential intrinsic conservation value of small land parcels.

It is emphasized that the *wrscr* measure does not take into account any complementarity or connectivity considerations and use of this measure does not replace a full Zonation analysis. Two areas could have equally high *wrscr*-values, but due to the occurrence of a completely different set of species, which is accounted for in a Zonation analysis but not by the *wrscr* measure. Note also that distribution smoothing and interactions influence *wrscr*-values, as it is calculated from the data that is used in Zonation computations, and this data includes effects of all transforms done to input maps.

Useful information

.rank.asc, .prop.asc and **.wrscr.asc** files are raster files, which can be used to produce map images in GIS software. For example, in ArcView choose "Import Data Source..." from "File" -options and select ASCII raster as your import file type. Select the file you wish to import, rename it and **do NOT choose integer cell values**.
The rank-file is the one that can be loaded with the -l -command, and that can be used in landscape comparison (see section 3.6.3).

.txt file
This file is written during the run. It keeps track of the files, settings and analyses used during the runs and is useful for tracing errors and checking that everything happened as supposed to. Note that some error messages or warnings may appear here. The content of the memo should be checked after a serious analysis run, to verify that correct options appear to have been used and that there are no worrisome error messages or warnings.

.run_info.txt file

This file will be created only after you have closed the program. You can use it to go back to see what happened in your analyses. The content of the .run_info.txt file is identical to that of the .txt file.

3.4.2 Optional output file formats

In addition to the output files Zonation produces automatically, it is possible to output the raster grid output in several different raster and picture formats. This is useful when you want to format your map outputs for publication purposes or further GIS analyses.

Image output

Zonation creates the color map with cell ranking in .jpg and .emf formats by default. In addition, it is possible to output the map in .png and .bmp formats. To output the additional image formats, type an additional parameter to your command line calling Zonation (see section 3.2.1. for details):

```
--image-output-formats [png] [bmp] [jpg] [emf]
```

Grid output

In addition to the ASCII rasters Zonation produces by default, you can output the raster grids in GeoTIFF (.tif) and Erdas Imagine file (.img) formats and compressed versions of both. All output formats except ASCII have 32 bit float element type. Tif produces a GeoTIFF file and img an Erdas Imagine file. Compressed GeoTIFF uses DEFLATE compression. Specify the output raster formats by typing an additional parameter in your Zonation call (see section 3.2.1. for details):

```
--grid-output-formats [asc] [tif] [img] [compressed-tif] [compressed-img]
```

3.4.3 Output files from optional analyses

.grp_curves.txt

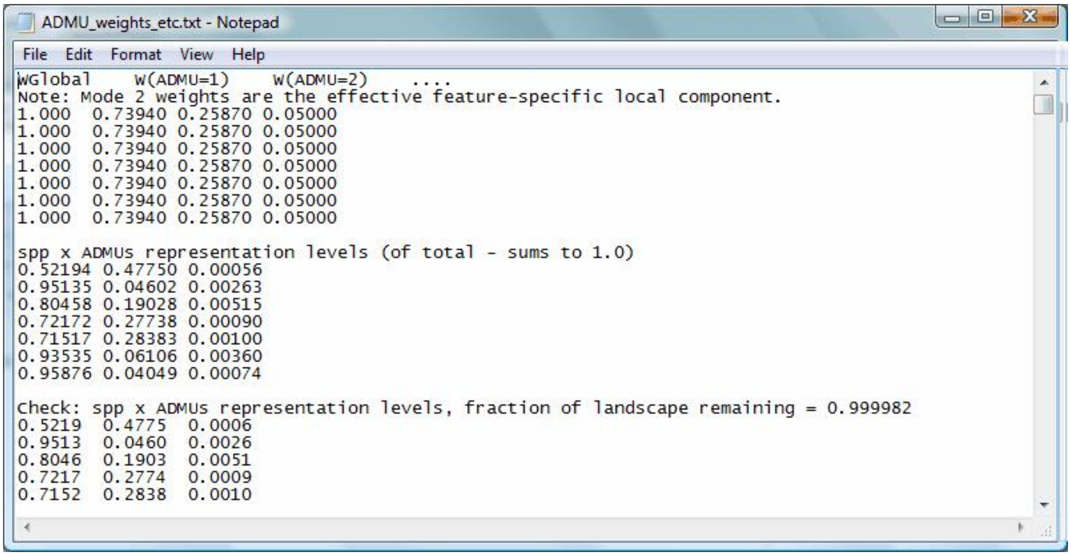
If you have assigned your species or other biodiversity features to output groups, Zonation will output an additional output file for you - a .grp_curves.txt file. This file contains representation curves for minimum, mean, and maximum representation in the course of cell removal for each group. To get this output, you need to include a groups file in to your analysis (see section 3.3.3.12. for details and settings). Output groups are specified in the first column of the groups file. Each species or other biodiversity feature can be linked to an output group.

Groups could be separately assigned, for example, to
(i) different higher taxa, birds, mammals etc
(ii) community and species features separately
(iii) negatively and positively weighted features in separate groups
(iv) habitat quality and connectivity layers separately
(v) etc., and in combinations of the above

All this information could be computed from the .curves.txt file, but that requires manual operations etc. which can be tedious. Thus, the grouped output is meant for saving manual work.

ADMU_weights_etc.txt

This output file is produced by the administrative units analysis. The file includes a matrix describing the joint effective weights for each administrative unit and species, as combined from global and local priorities and the local weights assigned a priori. The file describes species representation levels in each administrative unit as cell removal proceeds.



```
wGlobal     w(ADMU=1)     w(ADMU=2)     ....
Note: Mode 2 weights are the effective feature-specific local component.
1.000  0.73940 0.25870 0.05000
1.000  0.73940 0.25870 0.05000
1.000  0.73940 0.25870 0.05000
1.000  0.73940 0.25870 0.05000
1.000  0.73940 0.25870 0.05000
1.000  0.73940 0.25870 0.05000
1.000  0.73940 0.25870 0.05000

spp x ADMUs representation levels (of total - sums to 1.0)
0.52194 0.47750 0.00056
0.95135 0.04602 0.00263
0.80458 0.19028 0.00515
0.72172 0.27738 0.00090
0.71517 0.28383 0.00100
0.93535 0.06106 0.00360
0.95876 0.04049 0.00074

Check: spp x ADMUs representation levels, fraction of landscape remaining = 0.999982
0.5219  0.4775  0.0006
0.9513  0.0460  0.0026
0.8046  0.1903  0.0051
0.7217  0.2774  0.0009
0.7152  0.2838  0.0010
```

An example of the additional output from the administrative units analysis. Zonation first prints a species x weights matrix, where species are in rows and weights (global weight and local weights for administrative units 1, 2 and 3) in columns. The second matrix describes the representation levels in each administrative unit before any cells have been removed. Here, species are in rows and administrative units in columns. The following matrices describe the development in representation levels as cell removal proceeds. A matrix is produced every time after 1% of cells have been removed.

3.5 Post-processing analyses & options

This chapter includes descriptions for the three different types of analyses, that can be conducted for solutions produced from the main Zonation runs. These three groups are:

1. Post-processing analyses, which can be done manually from the Zonation windows interface. These include
   - Landscape identification
   - Statistics for management landscapes
   - Solution comparison
   - Fragmentation uncertainty analysis

2. Solution cross-comparison using loaded solutions.

These post-processing analyses are analyses that Zonation can perform for you automatically after the priority ranking has been computed. Some of these analyses could probably be done in GIS, albeit with significant effort.

3.5.1 Automated post-processing

After computing several Zonation priority ranking solutions, one may wish to examine the properties of a putative solution. With Zonation v.2.0, it was possible to do such post-processing manually. In Zonation v.3.0, post-processing analyses have been automated. Two types of automated post-processing analyses are available in Zonation: identifying management landscapes and comparing the spatial overlap between two solutions. These options are described in more detail in the following subsections.

3.5.1.1 Landscape identification

Landscape identification options allow identification of separate management landscapes based on the distance and similarity in species composition between two sites. Spatially distinct areas (consisting of multiple grid cells) in a Zonation solution can be classified into management landscapes. An area is joined to a landscape if it is close enough and similar enough in the species composition to *any* other distinct area in the same landscape. Landscape identification is done for a given fraction of the landscape.

Management landscapes can be identified among
- all cells in a certain top fraction of the solution
- cells that are masked into the top fraction of a solution
- a certain top fraction within cells that are masked into the top fraction of a solution.

Running landscape identification

To identify management landscapes from your Zonation solution, you need an automated post-processing file describing the fraction of landscape to be included, maximum distance and minimum similarity of cells to be included in the same management landscape, along with the possible mask file to be used. Please see section 3.3.3.17. for details of the input file contents.

Adjust your run settings file (section 3.3.2.3.) to indicate that you want to run automated post-processing. For this, you only need to type the name of your post-processing file:

```
post-processing list file = ppa_list_file_name.txt
```

The landscape identification analysis produces two output files:

nwout.#.ras.asc file	Here the matrix indicates which cells belong to which management landscapes. Each landscape has an integer starting from number one. If a cell has a value of -2, it means that the respective cells has not been included in the given top fraction (see "Percentage of landscape" above). Remember that this file, as any of the ASCII files produced with Zonation, can be imported to GIS programs. However, when importing this file, **select integers** as the format of you cell values!
nwout.#.spp_data.txt file	A text file containing statistics for species or other biodiversity features in management landscapes. See section 3.5.1.2. for contents of this file.

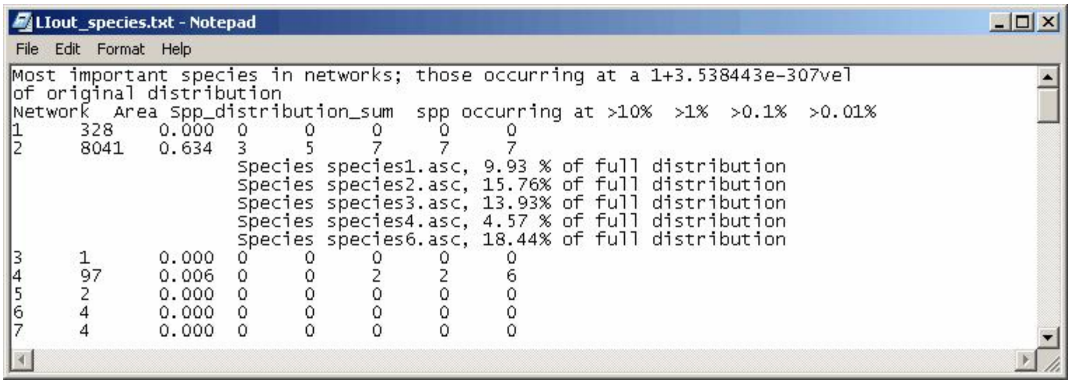
The number # in the file name is generated by the ordinal number of the call to an LSI analysis; the first call produces .nwout.1.ras.asc, the second .nwout.2.ras.asc and so on. The beginning of the LSI output files will be the general output file name you have given in the Zonation call.

Literature

The landscape identification analysis is described by
Moilanen, A., Franco, A. M. A., Early, R., Fox, R., Wintle, B., and Thomas, C.D. 2005. Prioritising multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society of London, Series B, Biological Sciences* 272: 1885-1891.

3.5.1.2 Statistics for management landscapes

The landscape identification procedure also produces a text output file (network species data) containing statistical information of the management landscapes.



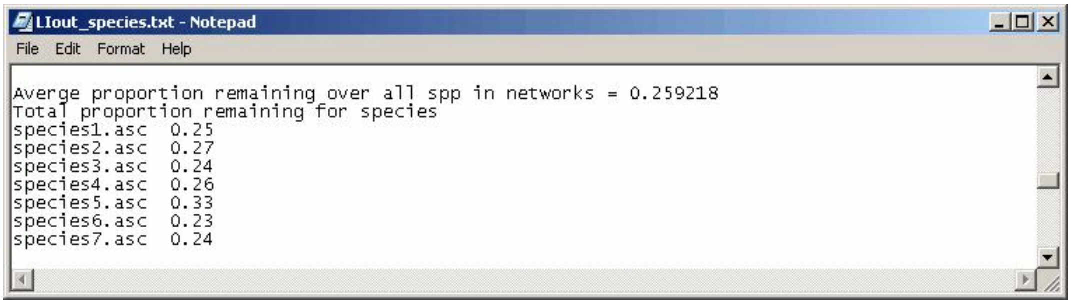
```
Most important species in networks; those occurring at a 1+3.538443e-307vel
of original distribution
Network  Area Spp_distribution_sum  spp occurring at >10%  >1%  >0.1%  >0.01%
1     328   0.000  0     0     0     0     0
2     8041  0.634  3     5     7     7     7
                    Species species1.asc, 9.93 % of full distribution
                    Species species2.asc, 15.76% of full distribution
                    Species species3.asc, 13.93% of full distribution
                    Species species4.asc, 4.57 % of full distribution
                    Species species6.asc, 18.44% of full distribution
3     1     0.000  0     0     0     0     0
4     97    0.006  0     0     2     2     6
5     2     0.000  0     0     0     0     0
6     4     0.000  0     0     0     0     0
7     4     0.000  0     0     0     0     0
```

With the file you can receive a set of information about the different management landscapes in your data. **The first part** of the file contains statistics about species occurrences in each landscape. This part is divided to eight columns:

1. Number of the management landscape.
2. Area of the management landscape in cells.
3. Sum of species distribution proportions. In other words this value shows how large proportions of species **original** distributions the respective management landscape covers.
4. Number of species which have more than 10% of their original distribution located in the management landscape.
5. Number of species which have more than 1% of their original distribution located in the management landscape.
6. Number of species which have more than 0.1% of their original distribution located in the management landscape.
7. Number of species which have more than 0.01% of their original distribution located in the management landscape.
8. Number of species which have more than 0.001% of their original distribution located in the management landscape.

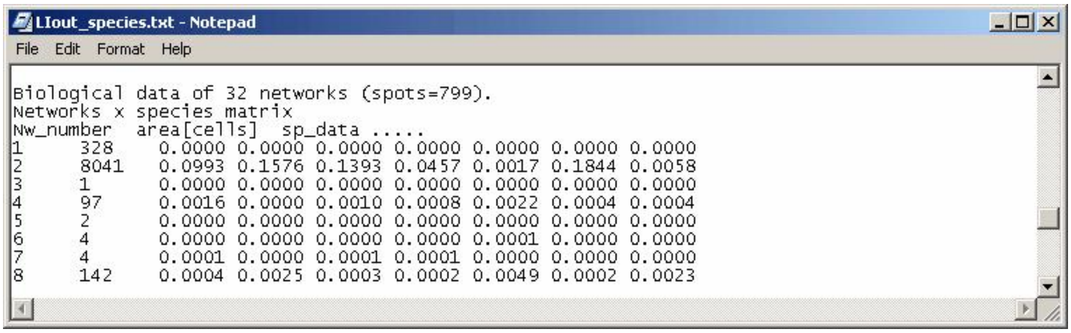
If the five last columns are marked as zero, it means that in that particular management landscape the species occurrence is less than 0.01% of their full distribution. Note that if any of the species have a larger proportion than 1% of its distribution located in the landscape, the program automatically prints a list of those species and the precise proportions of their distributions on the next rows.

The second part shows you how large proportions of species distributions are remaining in the whole landscape (all management landscapes together), that was initially included in to the analysis (the percentage of landscape defined in column 1 in your LSI command; section 3.3.3.17.). The program also automatically calculates an average of these proportions.



```
Averge proportion remaining over all spp in networks = 0.259218
Total proportion remaining for species
species1.asc  0.25
species2.asc  0.27
species3.asc  0.24
species4.asc  0.26
species5.asc  0.33
species6.asc  0.23
species7.asc  0.24
```

The third part contains a list of all management landscapes, area in number of cells and the proportions of distributions for each species in the respective management landscape. The species are listed here in the same order as they are in your biodiversity feature list file.



```
Biological data of 32 networks (spots=799).
Networks x species matrix
Nw_number  area[cells]  sp_data .....
1      328     0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
2      8041    0.0993 0.1576 0.1393 0.0457 0.0017 0.1844 0.0058
3      1       0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
4      97      0.0016 0.0000 0.0010 0.0008 0.0022 0.0004 0.0004
5      2       0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
6      4       0.0000 0.0000 0.0000 0.0000 0.0001 0.0000 0.0000
7      4       0.0001 0.0000 0.0001 0.0001 0.0000 0.0000 0.0000
8      142     0.0004 0.0025 0.0003 0.0002 0.0049 0.0002 0.0023
```

Note that setting the maximum distance between cells to zero allows you to view the statistics of every single spatially distinct patch in the landscape, but also increases the running time. A larger maximum distance leads to fewer management landscapes. Note that the program presently only allows identification of up to 30 000 landscapes.

3.5.1.3 Solution comparison

Solution comparison calculates how much two solutions overlap with each other and what is the average difference in the cell removal order. The comparison is always made between the present solution and a previously computed solution by using the rank.asc -files of both solutions as input files.

Running solution comparison

To find out how much two solutions overlap, you need an automated post-processing file (section 3.3.3.17.) with a line that has the following information:

Column 1: type "`LSC`" to indicate that you want to compare two Zonation solutions.
Column 2: top fraction of the **new solution** to be included in the comparison (value between 0 and 1)
Column 3: top fraction of the **previously calculated solution** to be included in the comparison (value between 0 and 1)
Column 4: file name of the previously calculated solution ranking map to be loaded, e.g. `my_previous_solution.rank.asc`
Column 5: file name for the output map showing the top fractions included for both solutions and the overlap between them, e.g. `overlap_between_1and2.asc`

Adjust your run settings file (section 3.3.2.3.) to indicate that you want to run automated post-processing. For this, you only need to type the name of your post-processing file:

```
post-processing list file = ppa_list_file_name.txt
```

Output

.ras.asc -file The values in the matrix are as follows:

0. The cell is not included in the top fraction
1. The cell is included in both solutions (overlapping areas marked as yellow)
2. The cell is included only in the present solution (light green areas)
3. The cell is included only in the older solution (dark green areas)

Remember that this file, as any of the ASCII files produced with Zonation, can be imported to GIS programs. However, when importing this file, select integers as the format of your cell values!

3.5.2 Solution cross-comparison using solution loading

This analysis is used, for example, in figure 4 of Moilanen & Wintle (2007) *Conservation Biology*, **21**:355-364, where effects of BQP on solutions originally calculated with the BLP are evaluated.

This is a major analysis, which can produce very important information, for example
- How well a solution produced without connectivity criteria works if connectivity is actually needed
- How much (apparent) conservation value is lost if a solution is developed requiring connectivity, which actually is not needed
- Likewise for the inclusion/exclusion of interactions between species distributions
- Surrogacy analysis; a solution developed for one set of species can be evaluated for performance across a completely different set of species.
- To obtain representation curves for the original, unexpanded community types in a community level analysis (section 5.3.1.)

Overall, the main point is that a solution can be developed using one set of criteria but post-hoc evaluated using another set of criteria.

When loading an old solution the program does not just display the solution, but removes cells from the landscape based on the ranking order of the old solution. Thus it is possible to test the performance of one network with different settings, using new settings to evaluate the solution. You can, for example, run the basic Zonation and then test how well the resulted network would perform, if uncertainty or boundary quality penalty would be included in the analysis. Full output will be produced from the loaded analysis, but cell removal order will be enforced according to the.rank.asc file that is loaded.

To do this, you need to load your old solution with new settings (see section 3.2.3. for instructions). Type "`-lfilename`" as the second parameter of your call. Remember also to enter the correct name of your adjusted settings file as the third parameter in your call.

You can compare the curves.txt files of the two solutions which reveal how large proportion of distribution of each species is remaining when landscape is iteratively removed.

Tutorial examples of solution loading are provided in exercises 6 and 10.

3.6 What Zonation does NOT do directly

Polygon data

Zonation does not operate on GIS data in polygon vector format. Such data needs to be converted into raster format before it can be used as Zonation input. Flexibility in the raster type and increased memory capacity alleviate the data intensity of raster files.

Multi-action planning

Zonation does not directly select from multiple alternative conservation actions. You can generate scenarios of where different actions could be applied, and then you can use Zonation for prioritizing among those spatially fixed actions. (See Thomson et al. 2009; Moilanen et al. 2011 CONL).

Dynamics

Zonation is not a dynamic modeling environment. One can fake dynamics by entering predictions for several time steps into the analysis (see e.g. Carroll et al. 2010; Thomson et al. 2009). In any case, Zonation does not model stochastic landscape or species dynamics, it is not a spatial PVA.

Corridors

Currently Zonation does not identify corridors that would maximize connectivity between areas in a path-like manner. (We are working on this...) Corridors might be developed using external software and entered as highly weighted feature layers into Zonation.

Habitat modelling

Zonation is not a statistical habitat modeling tool. Rather, Zonation uses output of habitat modeling as input. Zonation processes pattern of biodiversity occurrence after distribution modelling has first been done.

Fixing inadequate data

Zonation does not fix your data. Crap in, crap out. Zonation will produce a priority ranking no matter what you feed in. But, the value of the ranking depends on the quality of inputs. Z does not fix your data.

3.7 Data limitations & system requirements

Zonation v.3.0 has been developed for both 32- and 64-bit operating systems. The 32-bit software can use approximately 2GB of memory. The 64-bit version can use 4GB*4GB, which is effectively unlimited in principle. In reality, it is limited by what you can buy; presently a desktop PC running Windows 7 and having 32GB of memory does not cost much. High-end workstations have 200-300GB of memory. Increased memory capability allows the use of much more data than with v.2.0. Memory can be used for compensating for intelligence. More complicated setups and more feature layers are allowed with increased memory.

As the memory capacity of Zonation v.3.0 is increased massively, the software now becomes limited by speed for large data sets. To alleviate this, the Zonation core has been made partially multi-threading, which multiplies computation speed by a small integer factor. For example, data that has 28 million elements and about 50 layers runs in two days, if warp factor is set to 500. The practical limits for data set size thus are around 50 million effective elements, which corresponds to the world surface area at 2 km resolution.

Suitable organization of data and analysis settings allow for initially unlikely analyses to be possible. For more involved examples see:

Thomson et al. 2009 about habitat restoration.
Carroll et al. 2010 about climate change.
Gordon et al. 2009 about urban planning.
Lehtomäki et al. 2009 about extensions of forest conservation areas with relatively complicated connectivity arrangements.
Rayfield et al. 2009, which utilizes several different connectivity components per species.
(For full references, see section 2.1.)

Another thing that can be done with increased memory is faking of landscape dynamics. Enter layer sets for now and for several time steps in the future. This fakes dynamic landscapes and requires solutions that are balanced at all time steps, but is really only possible, with the increased memory capacity of Zv3.

Hardcoded limitations, which may be alleviated in later Zonation versions

All analyses

Maximum number of species (any biodiversity features)	25 000
Maximum SSI species (point occurrence lists)	25 000 (thus total maximum is 25 000 + 25 000)

Distribution smoothing

Maximum size of rasters	256 million grid cells

All of this can be effective data. It is worth noting, though, that when grid size reaches 30-50 million elements, computation time increases to days or weeks. This may be inconvenient.

Boundary Quality Penalty (BQP)

Maximum number of penalty curve profiles	50
Maximum number of points on a penalty curve	20
Maximum number of different species-specific buffer sizes	100

Landscape identification

Maximum number of management landscapes	30 000

Memory requirements

Memory requirements depend on the data you are using. Naturally, the larger the data (many species and/or high resolution and/or large landscape), the more memory you will need to run the computations. One raster file of one million elements (cells that have real data, not missing values) for one species requires 4 MB of memory. Thus you can roughly calculate the maximum number of species that you can use with the help of this formula:

Max. species = 0.7 * memory in MB / (4 * grid size [in millions of elements])

The 0.7 in the formula accounts for the memory needs of the operating system and the memory needs of Zonation in addition to the species data matrixes. Thus with 4 GB (= 4 000 MB) of memory you can have approximately (0.7 * 4 000)/(4 * 5) = 140 species with 5 mil. informative grid elements. Using BQP mode 2 (species-specific missing data areas) approximately triples the memory consumption compared to an analysis with no BQP used, leading to a respective loss in the number of species that can be sued in the analysis. Interaction layers count as independent features for the purpose of memory computations.

Note that this estimate of memory consumption is only indicative, but sufficient for getting an idea if an analysis definitely should or should not run. Check amount of available RAM in Windows task manager to verify that Zonation has not run out of memory.

3.8 Troubleshooting

Here is a short list of things to check, when encountering problems:

- **Directory paths**
  Check that you have entered the correct paths to your files, so that the program can find them. If you have the program in the same directory with your input files (recommended), you do not need to type the paths, just the file names. However, if some/all of your input files are located some where else, a directory path is needed for these files. If files are in a subdirectory from the exe-directory, then filenames can be entered as sub_dir_name\filename.

- **File names**
  One reason for problems can be the long directory and file names, especially if you are using the command prompt. Thus try to keep the directory names short (e.g. max. 8 characters). Also, do not use any spacings in your directory or file names. In Zv3 this is not a problem in principle, but avoiding special characters and spaces in file names is nevertheless safest.

- **Computer memory capacity**
  If the program is running very slow during computations, check Task Manager/Performance. If your RAM (Physical Memory) is close to zero, you have run out of memory. This does not mean that the program has jammed, but it will take a couple of lifetimes for it to finish up the calculations. In other words, stop the computation and try again after closing all other programs to save memory, or with a smaller data set or with a computer that has more memory. See section 3.8 for more detailed information related to computer memory capacity. It is not ok to have Zonation running using virtual memory (the hard drive), because that will simply run too slow.

- **Operating system**
  Zonation v. 3.0 is Windows 32bit or 64bit software, which should be operational for example on Windows XP, 2000 and 7. The 32 bit software runs on both 32 and 64 bit Windows versions; the 64bit Zonation only runs on 64 bit Windows.

- **Check the memo**
  Some warnings or error messages appear in the memo. Read through the text to check for any information that might give a clue to solving your problem.

- α **–values**
  Check the α –values for any errors. Remember that these values have to be in **same unit of length** as the cell size given in the species distribution map file. It is very easy to get these values wrong at first calculation - be sure to verify computations. A common error is that alpha is given per kilometer (alpha on the order of 1.0), but raster file units are in meters. In this case your alpha-values are 1000x too large, leading to a solution which is for practical purposes identical to the one you get when connectivity is not used at all.

- **Decimal points and commas**
  Always remember to use only decimal points in the settings files - this is required by the core! NO commas! Zonation assumes decimal dots, and commas will result is

undefined errors. If you come from a region that uses decimal commas, then beware potential for confusion. You can change the decimal setting from the Windows control panel.

- **Empty rows at the end of your input files**
  Check that you do not have any empty rows at the end of your input files. With bad luck this might cause some unexpected software behaviour.

- **Differences in grid sizes/cell sizes**
  All raster files should have the same grid size. This means that in all files the number of columns and rows as well as the size of cells should be equal. You can check these information from the beginning of each raster file.

- **No data -rows at the edges of species distribution matrixes**
  Computational efficiency requires the input data to have at least one row of no data on each edge of your species distribution grids. Otherwise the program will automatically transform all values on the edge rows to missing data. Usually this is not an issue, but can be so if your data goes right to the edge of the area.

- **Differences in grid matrixes between species distribution and other data files**
  Check that all species distribution rasters are congruent with any other raster files used in analysis (e.g. cost layer, uncertainty layers etc.) This means that all those cells in a grid which have data for any of the species (features) used in the analysis (that is to say, the cells that are NOT marked as "no data" in all species distribution files), also have to have a value in the optional raster grids. Equally, all cells marked as "no data" in all species distribution rasters should have the same definition in any optional rasters. Note that the analysis area can be masked using the analysis area mask.

Part

IV

4 Zonation Graphical User Interface

Zonation GUI (zig3gui) is a graphical user interface to the Zonation command line application (zig3). It allows you to manage analysis settings and input files and monitor the runs while they happen. Exporting the outputs and results in various formats is also possible with the GUI. This is a quick guide to the Zonation v.3.0 GUI. The GUI is under development - check the Zonation website http://www.helsinki.fi/bioscience/consplan/software/Zonation/index.html for updates!

Features

- Zonation project management
- Loading existing configuration and batch files
- View and edit input files
- Monitor Zonation calculations at runtime
- Edit and save output maps with custom colour schemes
- Compare maps between different analyses

Technical details

Zonation GUI is written in C++ using Qt toolkit by Nokia. It utilizes GDAL for GIS raster support. Various components of boost library are also used.
The interprocess communication between the GUI and the command line process is handled through shared memory mechanism.

4.1 Main menu

In the main menu of the Zonation v.3.0 GUI, four drop-down menus are available: **File, Tools, Window** and **Help**.

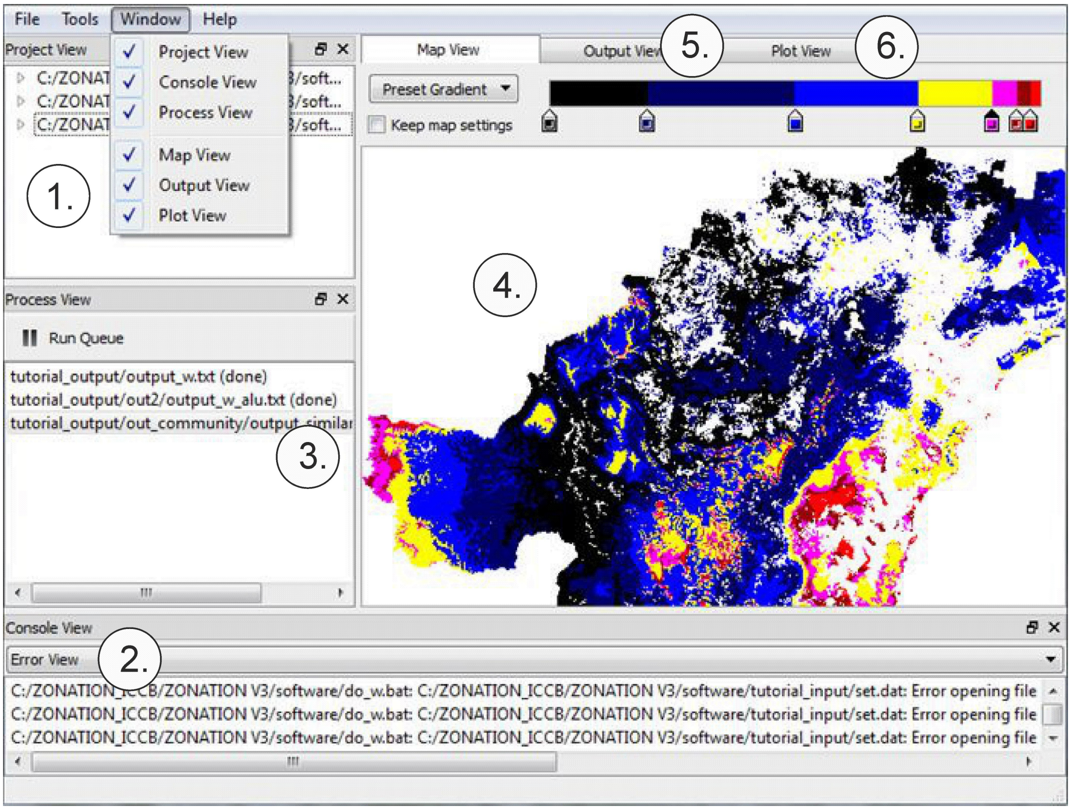
The File menu allows you to open a new project or view recent projects.

The Tools menu is for defining preferences.

The Help menu lets you access Zonation v.3.0 User Manual in a html format.

The Window menu allows you to determine which windows are visible. The GUI has 6 different windows:

1. **Project View** for calling, editing and visualizing input files (see section 4.2.).
2. **Console View** for tracking error messages and warnings during the runs.
3. **Process View** for monitoring the progress of analyses (see section 4.2.).
4. **Map View** for visualizing input and output maps (see section 4.3.1.).
5. **Output View** keeping track of all settings, input files and analyses included in one run (see section 4.3.2.).
6. **Plot View** visualizing the performance of biodiversity features during cell removal (see section 4.3.3.).


File
Tools
Window
Help
Project View
✓ Project View
✓ Console View
✓ Process View
✓ Map View
✓ Output View
✓ Plot View
C:/ZONAT
C:/ZONAT
C:/ZONAT
3/soft...
3/soft...
3/soft...
1.
Map View
Output View
5.
Plot View
6.
Preset Gradient
☐ Keep map settings
4.
Process View
Run Queue
3.
Console View
Error View
2.

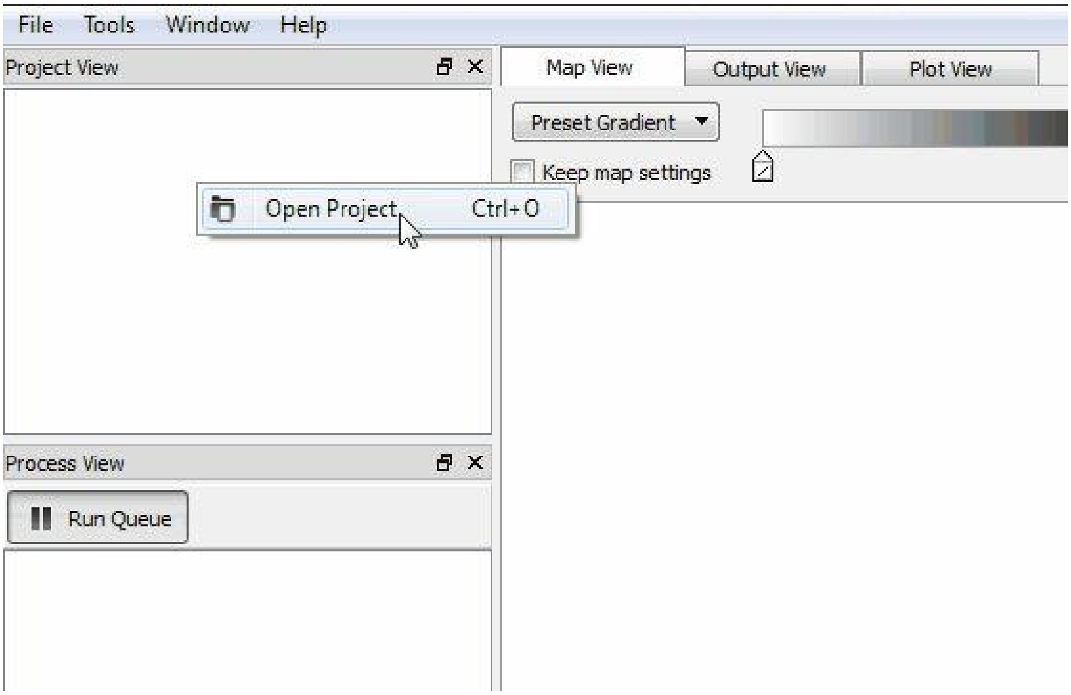
```
tutorial_output/output_w.txt (done)
tutorial_output/out2/output_w_alu.txt (done)
tutorial_output/out_community/output_similar
```

```
C:/ZONATION_ICCB/ZONATION V3/software/do_w.bat: C:/ZONATION_ICCB/ZONATION V3/software/tutorial_input/set.dat: Error opening file
C:/ZONATION_ICCB/ZONATION V3/software/do_w.bat: C:/ZONATION_ICCB/ZONATION V3/software/tutorial_input/set.dat: Error opening file
C:/ZONATION_ICCB/ZONATION V3/software/do_w.bat: C:/ZONATION_ICCB/ZONATION V3/software/tutorial_input/set.dat: Error opening file
```

4.2 Project management

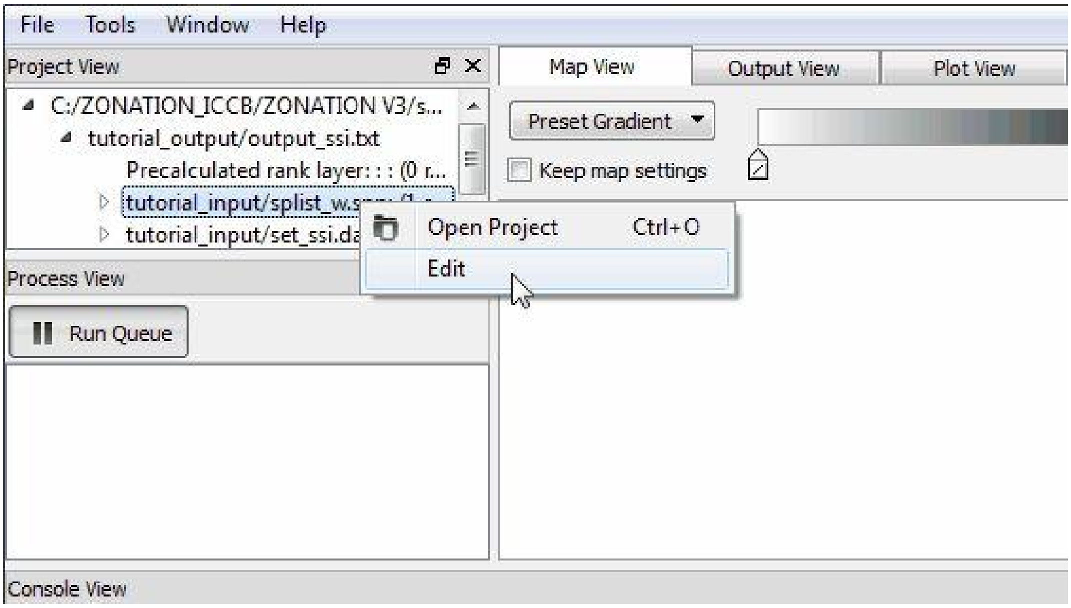
Open the **zig3gui.exe** file by double-clicking.

Project view is located in the upper left part of the GUI main window. By right-clicking your mouse on the white area, you can open a menu.
Select “**Open Project**”.

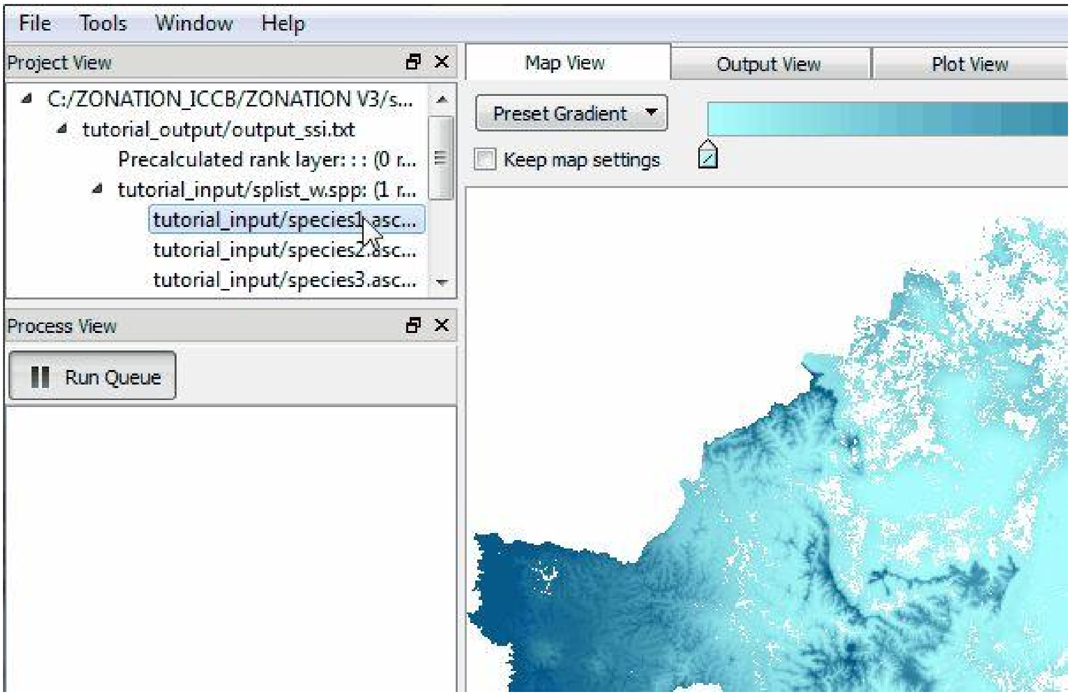

File
Tools
Window
Help
Project View
Map View
Output View
Plot View
Preset Gradient
Keep map settings
Open Project
Ctrl+O
Process View
Run Queue

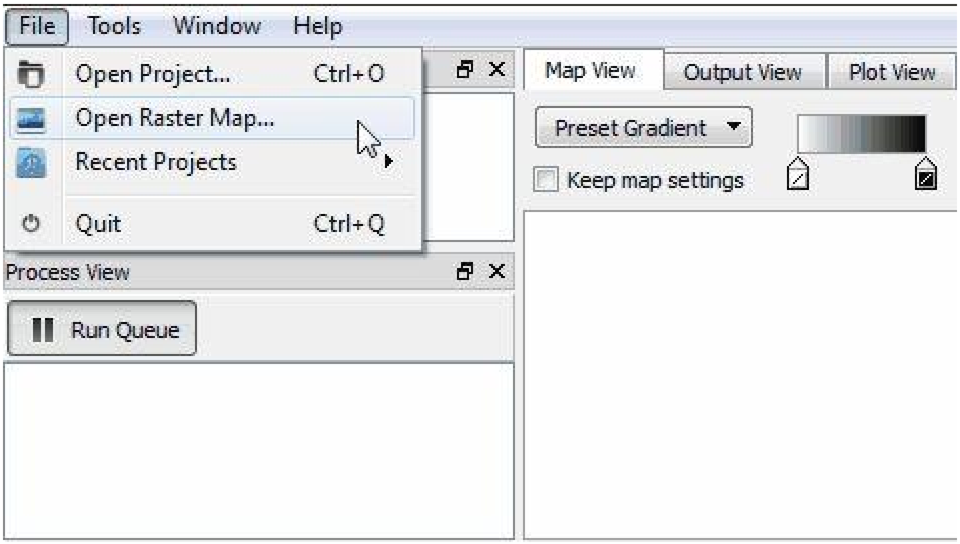
This allows you to browse for the right file. Select the batch file you want to use in the analysis. It should now appear in your Project window.

By clicking the small arrow on the left, you can view the file tree containing all input files which are referred to in the batch file or the subsequent file reference chain. By right-clicking the files and selecting "**Edit**", you can open the input files and edit them. The changes you make to the files will be saved into the files themselves.

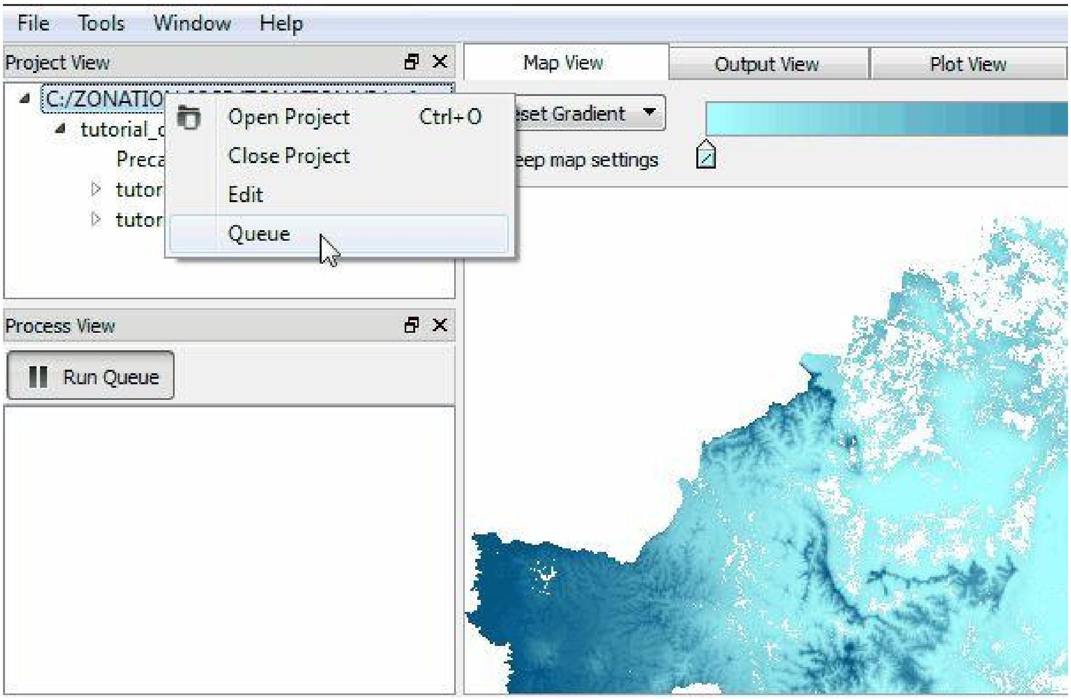

File
Tools
Window
Help
Project View
Map View
Output View
Plot View
C:/ZONATION_ICCB/ZONATION V3/s...
tutorial_output/output_ssi.txt
Precalculated rank layer: : : (0 r...
tutorial_input/splist_w.s...
tutorial_input/set_ssi.da...
Preset Gradient
Keep map settings
Open Project Ctrl+O
Edit
Process View
Run Queue
Console View

Input raster files can be viewed in the **Map View** (section 4.3.1.) by double-clicking the file name. Raster maps can be opened also from **File -> Open Raster Map**.

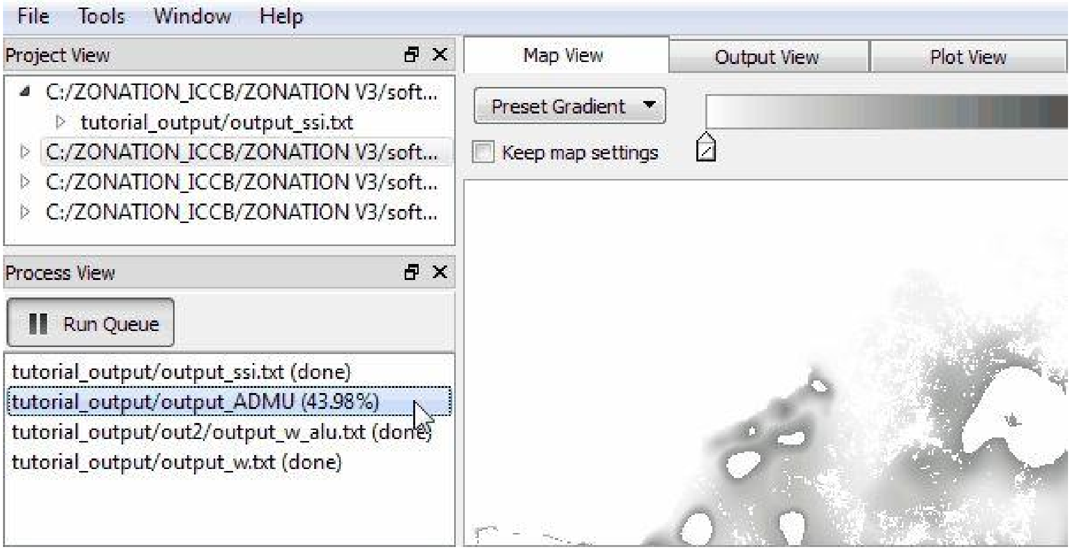

File
Tools
Window
Help
Project View
C:/ZONATION_ICCB/ZONATION V3/s...
tutorial_output/output_ssi.txt
Precalculated rank layer: : : (0 r...
tutorial_input/splist_w.spp: (1 r...
tutorial_input/species1.asc...
tutorial_input/species2.asc...
tutorial_input/species3.asc...
Process View
Run Queue
Map View
Output View
Plot View
Preset Gradient
Keep map settings


File
Tools
Window
Help
Open Project... Ctrl+O
Open Raster Map...
Recent Projects
Quit Ctrl+Q
Process View
Run Queue
Map View
Output View
Plot View
Preset Gradient
Keep map settings

The different command lines in the batch files are listed by their output file names. (Each line corresponds to one Zonation run.) It is possible to execute the command lines either all at once or one by one. Right-clicking either the batch file name or an individual command line in the Project View and selecting “**Queue**” adds the selected item to the work queue shown in Process View.


File
Tools
Window
Help
Project View
Map View
Output View
Plot View
C:/ZONATIO
tutorial_
Preca
tutor
tutor
Open Project
Ctrl+O
Close Project
Edit
Queue
set Gradient
eep map settings
Process View
Run Queue

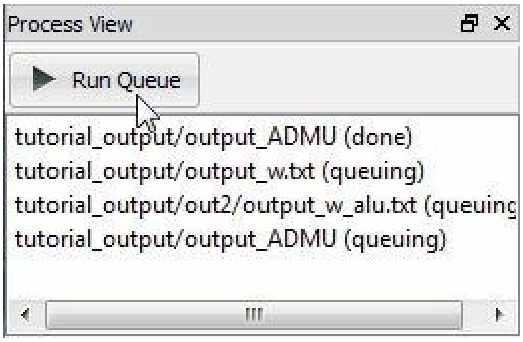
Your run will now appear in the **Process view** which is below the Project view. Double-clicking the line while Zonation is running will open the output map in the **Map view** (see section 4.3.). If your batch file has multiple lines, you can either run all of them at once by selecting the uppermost line or you can run individual command lines by adding them to queue.



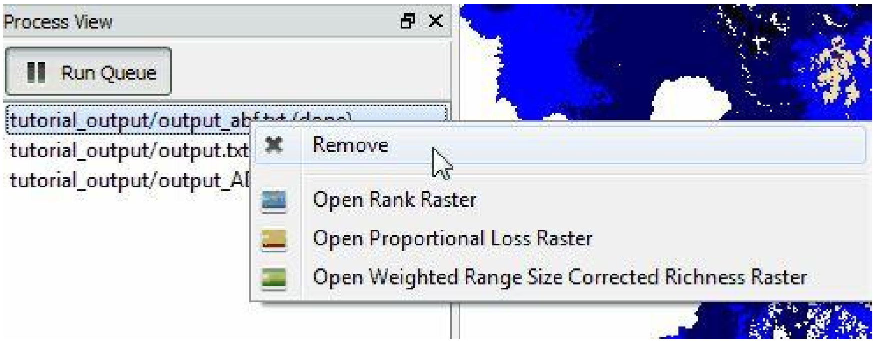
```
C:/ZONATION_ICCB/ZONATION V3/soft...
    tutorial_output/output_ssi.txt
C:/ZONATION_ICCB/ZONATION V3/soft...
C:/ZONATION_ICCB/ZONATION V3/soft...
C:/ZONATION_ICCB/ZONATION V3/soft...
```

```
tutorial_output/output_ssi.txt (done)
tutorial_output/output_ADMU (43.98%)
tutorial_output/out2/output_w_alu.txt (done)
tutorial_output/output_w.txt (done)
```

The queue can be started and stopped by clicking the **Run Queue**-button located at the top of the Process View. Pausing the queue will prevent further processes from opening but will not stop already running processes. Items can be removed from the work queue by right clicking them and selecting **Remove** from the opening popup menu. Removing a running process will terminate it forcefully. Right-clicking the run also allows you to display, in addition to the default rank raster, the proportion remaining and weighted range-size normalized richness layers. Note that when displayed, the .wrsnr.asc raster has been rescaled to the interval [0,1]. Thus, if looking at pixel-level information, the number that you see is "weighted range-size normalized richness relative to the maximum that occurs anywhere in the landscape".



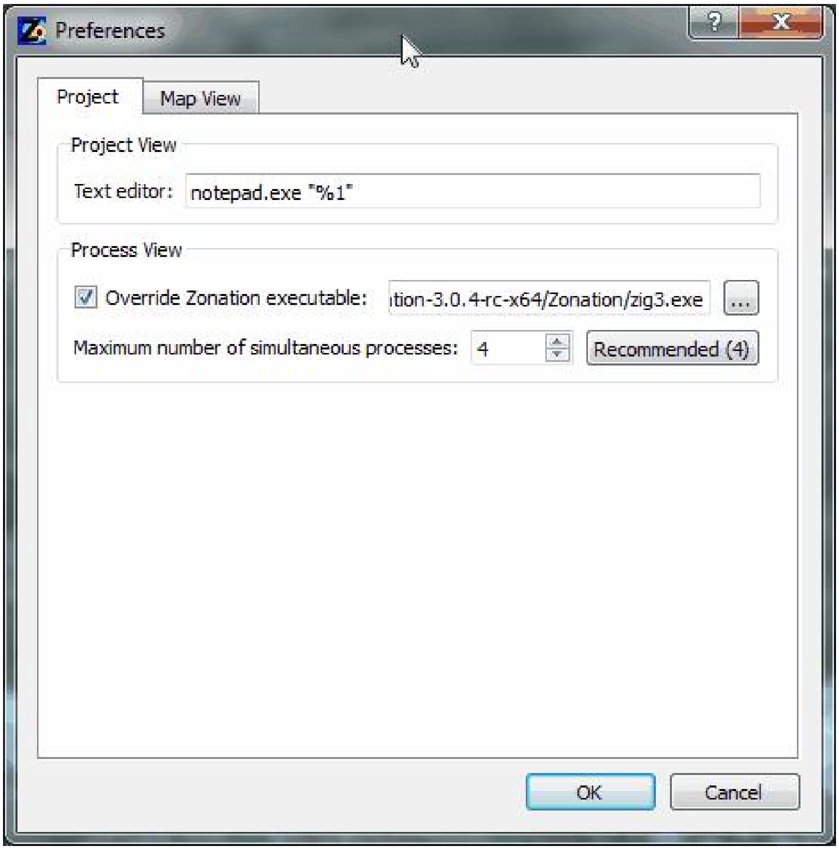
```
tutorial_output/output_ADMU (done)
tutorial_output/output_w.txt (queuing)
tutorial_output/out2/output_w_alu.txt (queuinc
tutorial_output/output_ADMU (queuing)
```


Process View
Run Queue
tutorial_output/output_abf.txt (done)
tutorial_output/output.txt
tutorial_output/output_AI
Remove
Open Rank Raster
Open Proportional Loss Raster
Open Weighted Range Size Corrected Richness Raster

By default, Zonation will call the core executable (zig3.exe) bundled with the GUI instead of the executable called in your batch file. This can be changed in the **Preferences** menu in **Tools-> Preferences**. The batch file can be any regular command prompt script. Please note however that the Zonation core executable called in the batch file must be reachable and must be of version 3.0.3 or higher. Otherwise the Zonation GUI will not be able to parse the batch file and your project will appear empty. Earlier versions (< 3.0.3) of Zonation have a different parser for batch files and are not interoperable with later versions.

Zonation GUI can run multiple items at once. By default, it won't try to run more processes at once than there are physical processors in your computer. The maximum amount of simultaneous processes can be controlled from the **Preferences** menu in **Tools-> Preferences**. In general it is most efficient to have several runs progressing simultaneously, but the ability to do so may be limited by the RAM memory needs of the runs and the memory available on your computer.

The memory issue: All of the simultaneous Zonation processes must fit into the physical RAM memory of your computer. (You can check memory usage from the Windows task manager.) Often in large zonation analyses it is the memory consumption that will be the bottleneck instead of the number of processors. If this is the case, lower your settings for the maximum number of processes so that all of the simultaneous instances can fit into memory. If the processes need to use virtual memory (= your hard drive), the computations will become so slow that they will never finish.


Preferences
?
Project
Map View
Project View
Process View
...
Recommended (4)
OK
Cancel

Text editor: notepad.exe "%1"

☑ Override Zonation executable: ition-3.0.4-rc-x64/Zonation/zig3.exe

Maximum number of simultaneous processes: 4

4.3 Visual output

Zonation v.3.0 GUI produces three types of visual output: the map from the cell ranking in **Map View** (section 4.1.1.), a memo describing the analysis in **Output View** (section 4.1.2.), and performance curves describing the representativeness of the solution in **Plot View** (section 4.1.3.).

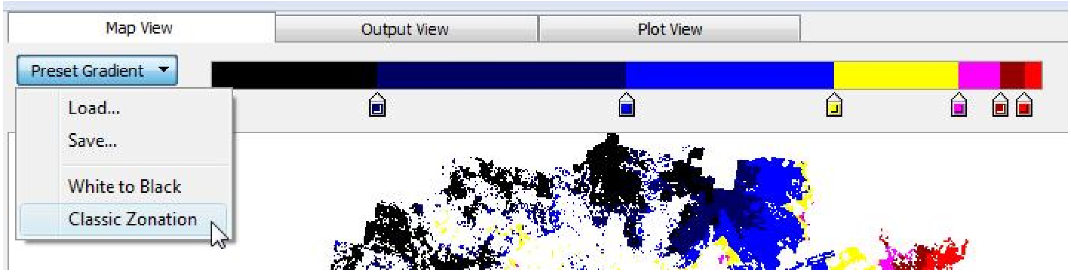
4.3.1 Map View

Zonation GUI shows the proceeding of cell ranking and the resulting map in the **Map view**. Double-clicking each run in the **Process view** (see section 4.2.) will open the respective output map in the Map view.

Zooming in and out is possible with the scroll wheel on your mouse. You can pan the map by dragging it with the mouse.

Colour schemes

The new Zonation GUI provides a tool for editing the output map to best suit your analysis and visual preferences. The default map has a white to black gradient, black being the highest priorities. In the **Preset Gradient** drop-down menu, you can choose to use either the **Classic Zonation** colour scheme or adjust the default **White to Black** gradient yourself.



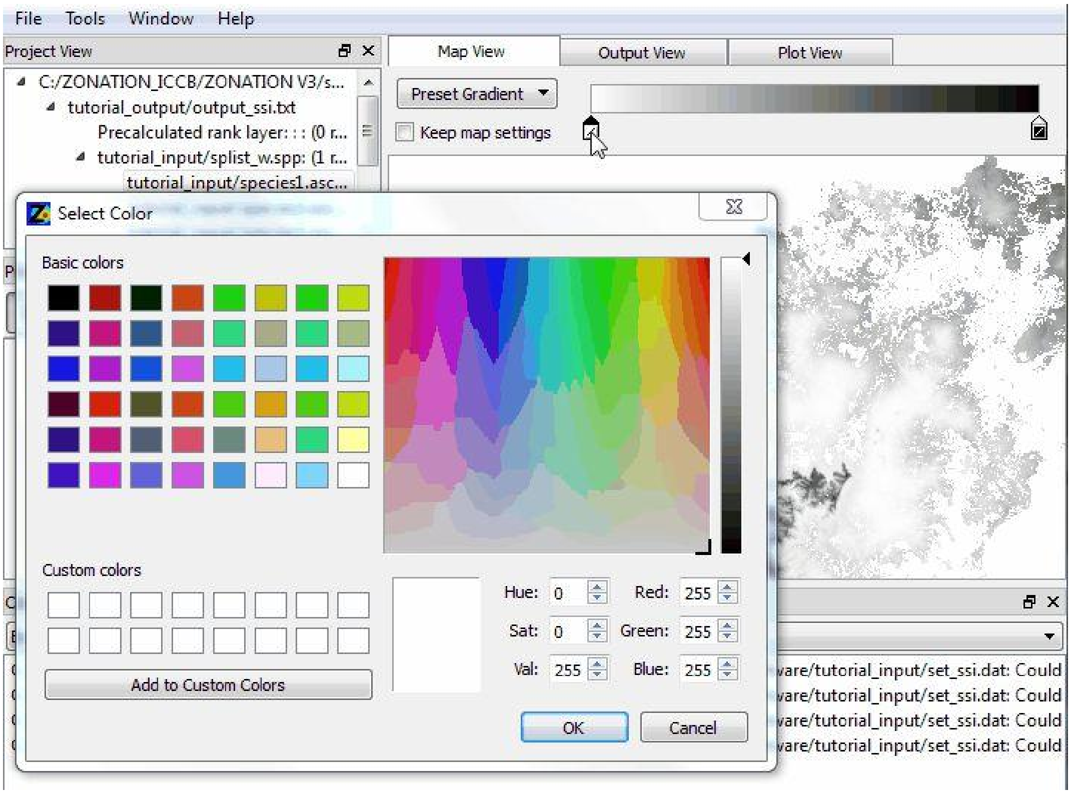
The Classic Zonation colour scheme shows the nested ranking on a map. The ranking of sites is visualized by using different colors to indicate the biological value of the site:

- red = the best 2 % of the landscape
- dark red = the best 2-5%
- magenta = the best 5-10%
- yellow = the best 10-25%
- light blue = the best 25-50%
- dark blue = the best 50-80%
- black = the best 80-100% (or the least valuable 20%)

The information of this map is equal to the .rank.asc file that the program produces as part of file output. This map will also automatically be saved as a picture (output.jpg and output.emf files), but you can save it again (e.g. with a different name or to a different directory) by right-clicking the picture. The background, i.e. the cells for which no data exists, are shown in white. In the beginning of analysis, before overwritten by the ranking, locations with SSI species are shown as red dots.

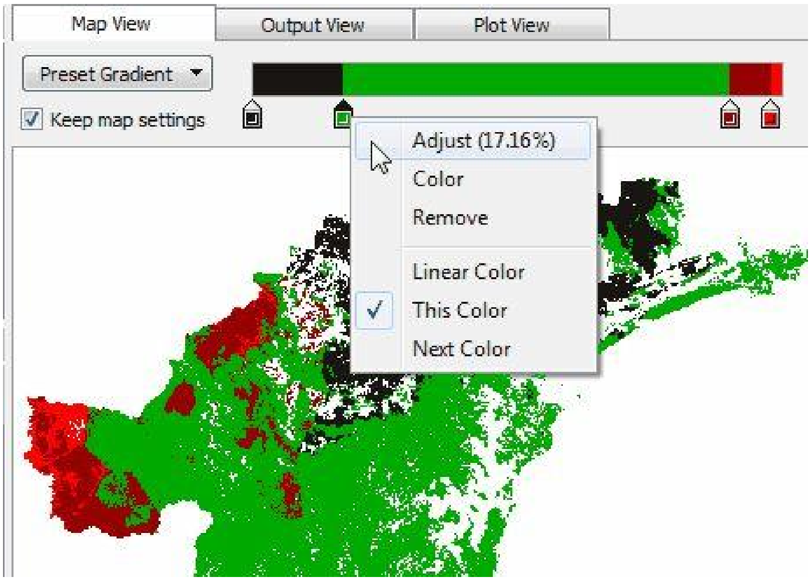
The new GUI has a **custom colour gradient** which you can design yourself. This gradient can be used for identifying any given top fraction of the landscape or any given bottom fraction of the landscape. It can also be used for identifying reserve network expansions in a hierarchic analysis. It can furthermore be used for identifying areas that satisfy a given minimum level for all species (features) - this can be achieved by coloring the .prop.asc output file. The custom color gradient tool corresponds to the rank and remaining buttons of the Zv2 GUI.

Right-clicking the mouse-pointer in the **Colour scheme** opens a menu from which you can adjust the colour tab settings. By selecting "**Adjust**" you can define the proportion of the priority ranking to which the gradient is applied. This is expressed in percentages 0-100%, with for example 99 indicating the top 1% point. Selecting "**Colour**" opens a new window through which you can adjust the colour you wish to assign to each colour tab. Right-clicking the grey space under the colour gradient adds a new cursor and opens a menu from which you can adjust the tab settings. Removing a cursor is possible by selecting "**Remove**" from the cursor menu. The appearance of the color scale is adjusted by assigning "**This color"**, "**Next color**" or "**Linear color**" to a color tab. Understanding how these work is perhaps easiest figured out by experimentation.

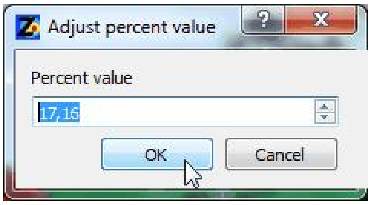

File
Tools
Window
Help
Project View
C:/ZONATION_ICCB/ZONATION V3/s...
tutorial_output/output_ssi.txt
Precalculated rank layer: : : (0 r...
tutorial_input/splist_w.spp: (1 r...
tutorial_input/species1.asc...
Map View
Output View
Plot View
Preset Gradient
Keep map settings
Select Color
Basic colors
Custom colors
Add to Custom Colors
OK
Cancel

☐ Keep map settings

Hue: 0
Sat: 0
Val: 255
Red: 255
Green: 255
Blue: 255


Map View
Output View
Plot View
Preset Gradient
Adjust (17.16%)
Color
Remove
Linear Color
This Color
Next Color

☑ Keep map settings


Adjust percent value
OK
Cancel

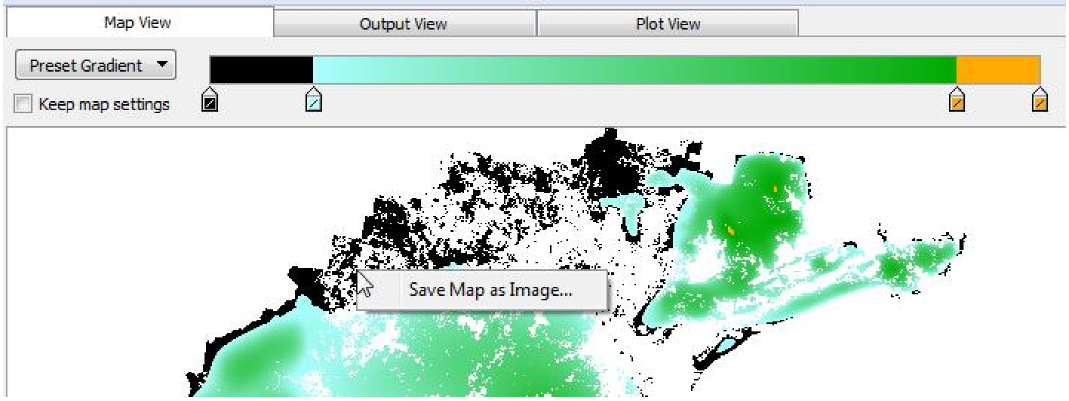
Percent value 17,16

You can save the colour scheme to use the same scheme later for other runs by selecting "**Save**" in the **Preset Gradient** menu. Previously saved colour schemes can be loaded by selecting "**Load**" and browsing for the file through the same menu.

Identifying a reserve network expansion. Let's assume your present reserve network is 7% of the landscape and that you are planning an 1.5% expansion. You first do a Zonation run where a hierarchy is forced in the analysis, with the existing reserves retained latest. You can then color the output of this analysis so that on color tab is placed at 93 (corresponding to the top 7%), and another at 91.5 (the 1.5% expansion). Give these tabs different colors to show the expansion clearly.

Saving output map with a custom colour scheme

During each run, Zonation automatically produces the ranking map in various picture formats (see section 3.4.1. for details). These maps follow the Classic Zonation colour scheme. You can save the output map with your custom colour scheme in a **.png** format by righ-clicking mouse on the map and selecting "**Save Map as Image**".

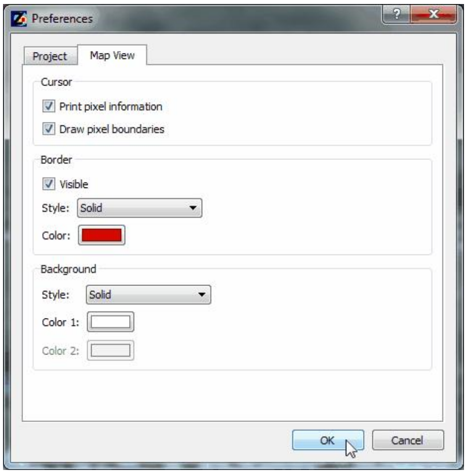


Visual comparison of outputs between runs

After you have run a number of runs, you can compare the outputs visually directly in the GUI. Double-clicking a line corresponding to a run in the **Process View** will open the Map View for that run. When switching between maps, it is often convenient to keep the zoom and pan settings as well as the colour scheme the same. You can do this by selecting " **Keep map settings"** in the Map view.

Map View Preferences

Zonation GUI has a useful pixel inspection tool which can be turned on from Tools -> Preferences menu. With this mode on, the value of the pixel under the cursor will be printed on screen. Preferences also contain settings for drawing the map border and background.


Preferences
Project
Map View
OK
Cancel

Cursor

☑ Print pixel information

☑ Draw pixel boundaries

Border

☑ Visible

Style: Solid

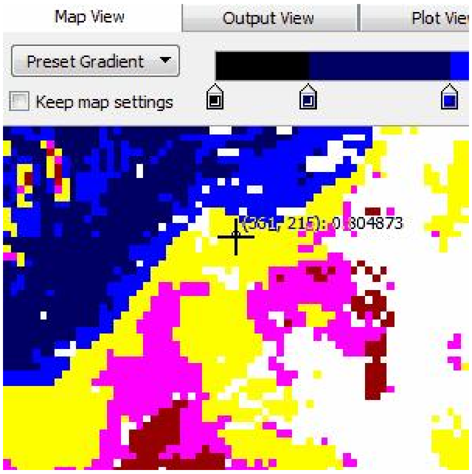
Color: 🟥

Background

Style: Solid

Color 1:

Color 2:


Map View
Output View
Plot Vie
Preset Gradient

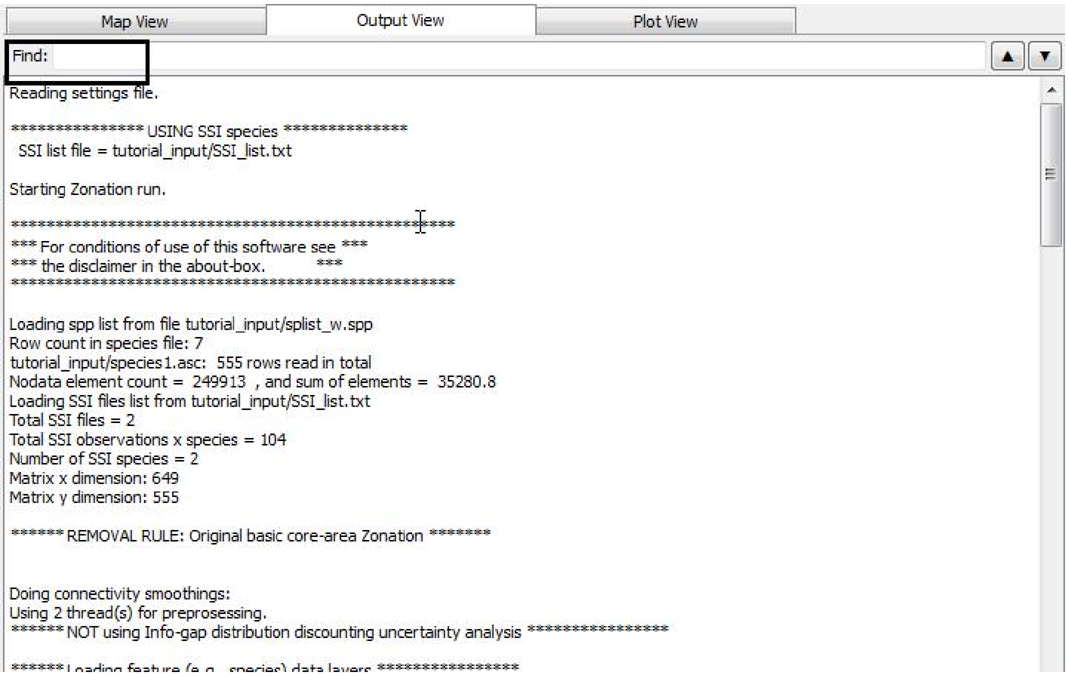
☐ Keep map settings

(361, 215): 0.304873

4.3.2 Output View

All calculations, settings and input files used during a Zonation run are stored in a memo file (section 2.4.2.). They are automatically also printed in separate window, the **Output View**, for each run.

Any error messages or warnings are printed in this window as well. It is advisable to go through the memo to check that no errors occurred during the analysis, and that you indeed used the correct input files and settings. The **Find** window, with which you can look for certain key words (we recommend 'error' and 'warning' and 'unable'), is located at the top of the Output View.



```
Reading settings file.

**************** USING SSI species ***************
  SSI list file = tutorial_input/SSI_list.txt

Starting Zonation run.

*****************************************************
*** For conditions of use of this software see ***
*** the disclaimer in the about-box.          ***
*****************************************************

Loading spp list from file tutorial_input/splist_w.spp
Row count in species file: 7
tutorial_input/species1.asc:  555 rows read in total
Nodata element count =  249913  , and sum of elements =  35280.8
Loading SSI files list from tutorial_input/SSI_list.txt
Total SSI files = 2
Total SSI observations x species = 104
Number of SSI species = 2
Matrix x dimension: 649
Matrix y dimension: 555

******* REMOVAL RULE: Original basic core-area Zonation ********


Doing connectivity smoothings:
Using 2 thread(s) for preprosessing.
******* NOT using Info-gap distribution discounting uncertainty analysis ******************

****** Loading feature (e.g. species) data layers *****************
```

4.3.3 Plot View

Curves visualizing the proportion of distributions remaining, average extinction risk, cost needed to achieve a given conservation value and proportions of distributions remaining are plotted in the **Plot View** for each run. The lowest fraction across species or other biodiversity features is plotted with a red line, whereas blue line represents the average across species or other features. When the Plot View is open in the GUI, you can view the curves from different analysis active in the GUI. Double-clicking a line corresponding to a run will change the Plot View to show curves from that run.

Four types of curves are plotted during each run:

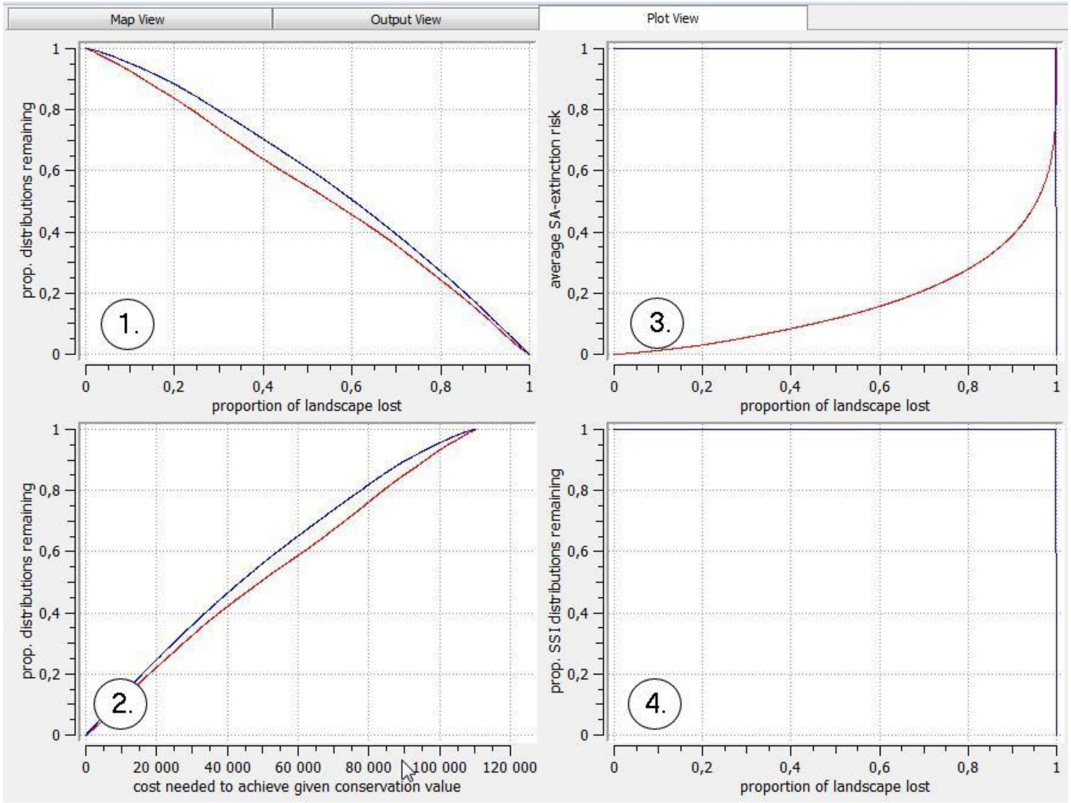
1. The first of the four plots shows summary information about the **proportion of distribution remaining** across species when landscape is removed. The red line represents the species with the lowest distribution remaining, and the blue line represents the average over all species. All this information can be found in numerical form from the **.curves.txt** file (section 3.4.1.), which also contains the respective curves for each species included in the analysis.

2. Below the distribution curve is the **cost curve** showing how high costs are needed for buying the respective top fraction. If no cost layer is used, all cells receive an equal value of 1.0 and the curves thus show the number of cells needed for respective top fractions. The curves begin from the proportion of landscape that is left after the initial removal of the worst proportion. If nothing is removed, the curve starts from zero, as in the picture below.

3. The third curve shows how the **extinction risk** of species increases as landscape is removed. This curve is based on the species-area ratio and shows the average extinction risk over all species, assuming the exponent *z* given in the settings.

4. The fourth and last curve displays the **proportion of distribution remaining for SSI species** (Species of Special Interest) when landscape is removed. Also here, the red line represents the species with the lowest distribution remaining, and the blue line represents the average over all species. Note that, unlike the other panels, this panel will only be displayed when SSI species are included into the analysis.

The information of these four graphs is equal to the .curves.txt file that the program produces as part of file output.


Map View
Output View
Plot View
1.
prop. distributions remaining
proportion of landscape lost
0
0,2
0,4
0,6
0,8
1
2.
prop. distributions remaining
cost needed to achieve given conservation value
20 000
40 000
60 000
80 000
100 000
120 000
3.
average SA-extinction risk
proportion of landscape lost
4.
prop. SSI distributions remaining
proportion of landscape lost

Part V

5 Zonation analysis setups for common planning needs

In the following sections we describe some analysis setups for Zonation that correspond to typical needs of managers and researchers. The idea is to facilitate the use of Zonation for practical needs, so that relevant parts of the documentation are identified and linked to, instead of having to read the entire manual. However, we recommend getting acquainted with the theoretical basis of the methods (section 2) and examples of analysis implementations from the literature. Some setups have corresponding tutorials with example files, but not all of them.

5.1 Basic analysis components

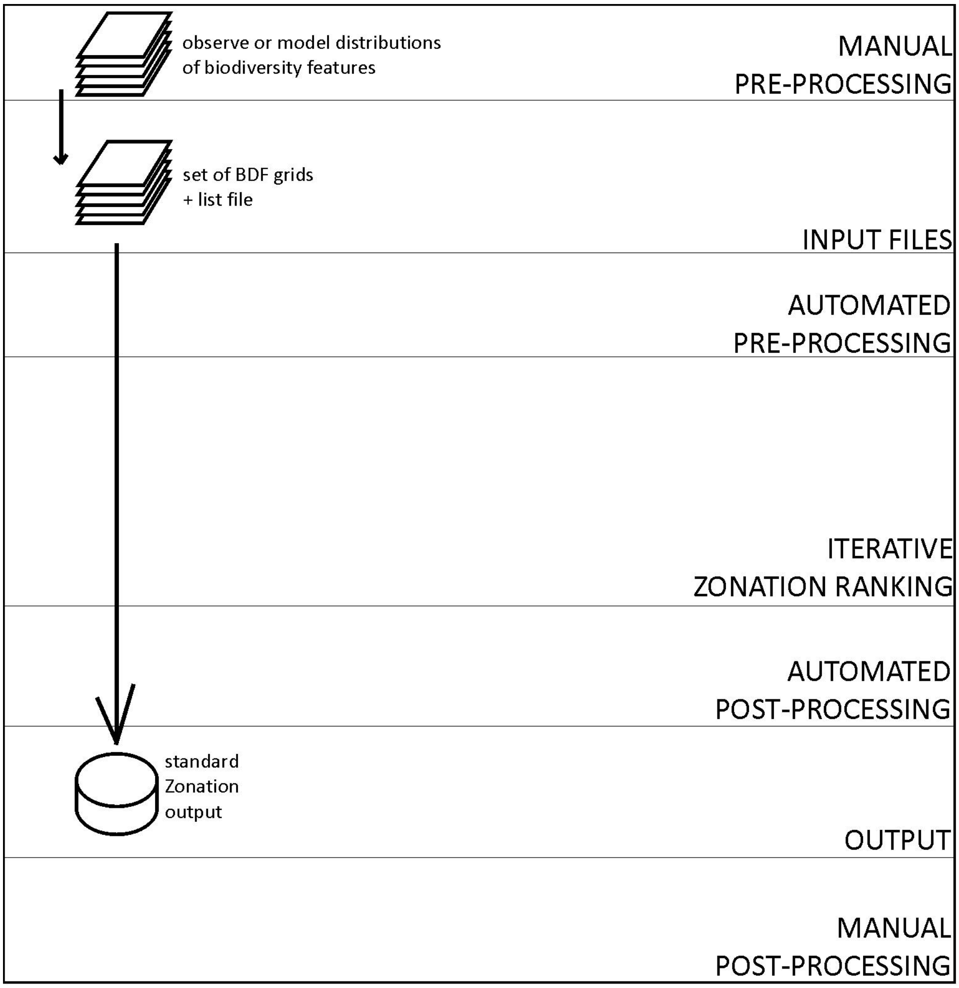
This section describes the use of basic components that are often included in Zonation prioritization. These components can be applied to any analysis (but please remember that including several connectivity considerations in a single analysis causes trouble for interpretation of the output and is not advisable).

5.1.1 Simple Zonation and species weighting

Planning problem to be solved

Simple non-spatial Zonation is frequently the first analysis one would run for any planning case. In most cases, this analysis is not intended to be a final product that would be used in the planning. Rather, non-spatial Zonation can be used to make sure all input files function and to get a general idea about what the output might look like. Simple Zonation can also be useful as a reference analysis. For example, one can compare this solution to another that includes connectivity considerations or other more complex settings. Representation curves of the solutions can be compared to assess, for example, how much local quality must be traded off for increased connectivity. As this analysis is simple and quick to run, it is often worthwhile to try different cell removal rules to see how they affect the outcome. Note that different assumptions about relative richness and rarity are implemented when conservation value is aggregated depending on whether you choose the core-area Zonation, additive benefit function or generalized benefit function as your cell removal rule. Computing the analysis using both core-area Zonation and additive benefit function can help you find out whether there are areas that support narrow-range species or other biodiversity features in an area with otherwise low richness in biodiversity features.

Process chart for the analysis


observe or model distributions of biodiversity features
MANUAL PRE-PROCESSING
set of BDF grids + list file
INPUT FILES
AUTOMATED PRE-PROCESSING
ITERATIVE ZONATION RANKING
AUTOMATED POST-PROCESSING
standard Zonation output
OUTPUT
MANUAL POST-PROCESSING

Examples from literature

Moilanen, A., and Wintle, B. A. 2007. The boundary-quality penalty: a quantitative method for approximating species responses to fragmentation in reserve selection. *Conservation Biology*, 21: 355-364.

The theory and algorithms behind basic Zonation is explained in sections 2.2. and 2.3.

Pre-processing of inputs

Processing of inputs often includes fitting habitat suitability models to existing species data and creating spatial predictions of habitat suitability or occurrences of biodiversity features across the planning region.

Input files

To run simple Zonation, you need

- A set of biodiversity feature grid layers (section 3.3.2.1.)
- A biodiversity feature list file (section 3.3.2.2.).
- A run settings file with appropriate settings (section 3.3.2.3.)

Analysis stages and settings

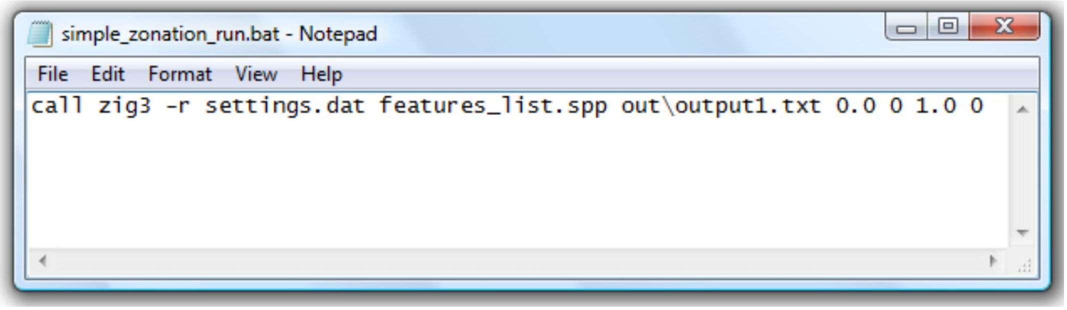
Here we describe the basic setups for running the software. The skeleton of the simple setup can and should be complemented with spatial and ecological considerations to add realism in the analysis. Descriptions for more complex setups are described in the next sections, but they always follow the same basic scheme.

To set up a Zonation prioritization run, you need to

1. Set all options of **additional analyses** (e.g. BQP, uncertainty analysis etc.) to zero in your **run settings file** AND in the **command line call** to indicate that no additional analyses are used.
2. Adjust your settings in the run settings file for the following options if necessary:
    - cell removal rule (see section 2.2.)
    - warp factor
    - edge removal
    - add edge points
    - logit space

   Note that if you do NOT select the edge removal, the computation times will increase significantly with large data sets.
3. Adjust the species weights in **species list file** if you wish to stress the conservation of certain species (e.g. rare species of high conservation value or commercially valuable species).
4. Type the call for Zonation in the command prompt and press enter to initiate the computation.

See section 3.2.1. for how to use batch files to call Zonation.



```
call zig3 -r settings.dat features_list.spp out\output1.txt 0.0 0 1.0 0
```

No special settings are needed for this analysis. Weighting of biodiversity features and choice of cell removal rule can affect the outcome.

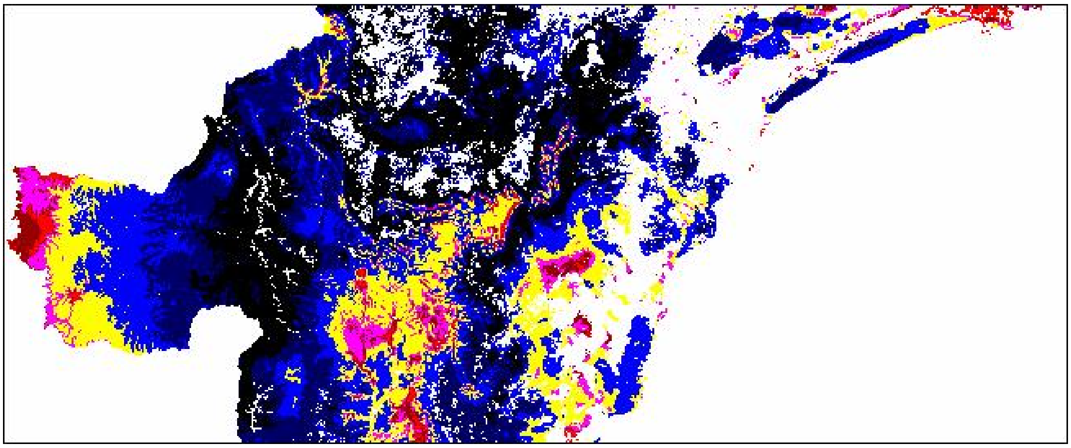
Weighting of species is a critical component of the algorithm. Problems associated with different initial sizes of species distributions are circumvented in Zonation by assigning a value for the full distribution of each species. By default, these values are equal, but species can be assigned differential weights (in the **species list file**) based on, for example, their taxonomic status, global rarity, economical value or population trend. Weighting of species affects the order in which cells are removed from the landscape. Cells that include a part of the distribution of a valuable species (high weight) remain later in the iterative cell removal process than cells only containing low-weight species, assuming everything else is equal between the occurrences. Weighting influences the fraction of a species distribution retained at any point of the cell removal. When using weighting, high-weight species retain a relatively higher proportion of their distribution. Also, note that the balance in representation levels developed by Zonation is such that narrow-range species typically have a larger fraction of their ranges protected compared to initially wide-ranging species.

Weights can also be used to test the efficiency of **surrogate species**. This is done by weighting the surrogate species normally (e.g. by 1) and giving a weight of 0 to all those species that are NOT used as surrogates. This way the program will not use the non-surrogate species in the selection of the next site to remove, but it will monitor the decrease of the distributions of these species as well. Thus, you can test how well a reserve network selected using surrogates will protect all species.

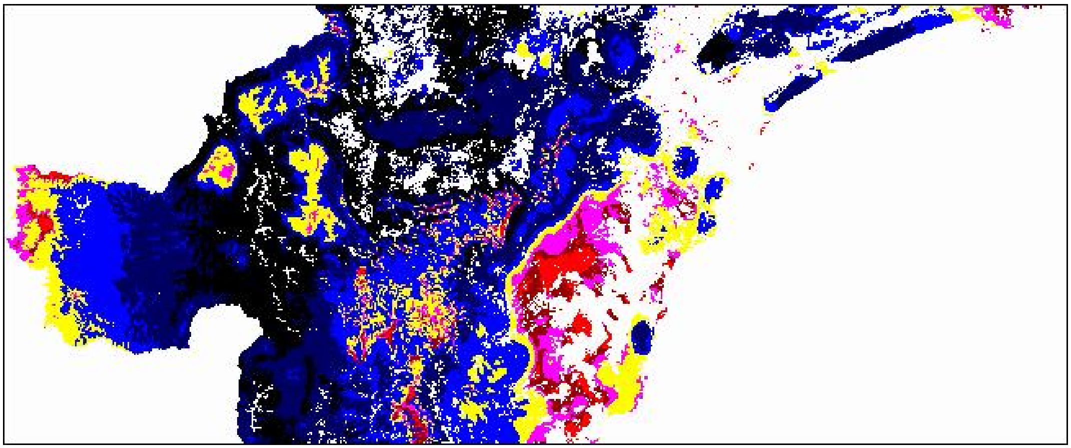
Output and its interpretation

Priority ranking map and representation curves can be used as reference for more complicated analyses.

Here is an example of how species weighting can influence the final solution. The pictures show the results of two basic Zonation runs for seven species. In the first picture no species weights have been used where as in the second picture one of the species has received a weight of 3.0.



Picture of solution when no weights are used.



Picture of solution when one of the species has a weight of 3.0.

Strengths, weaknesses and further considerations

Strength: Simple Zonation is simple and quick to set up and run.

Weakness: The analysis is too simple to be realistic. Considerations of connectivity, costs, landscape condition and retention and other such factors are missing, although they can have significant influence on the conservation outcome at the implementation stage.

Link to tutorial

See exercise 1 for a tutorial example of the basic use of Zonation. Exercise 2 provides an example of species weighting.

5.1.2 BLP Zonation

Planning problem to be solved

Accounting for connectivity in conservation planning. Boundary length penalty (BLP) is a simple method to induce aggregation to the protected area network. Increasing the total length of boundary will decrease conservation value of the remaining cells, and thus the algorithm will seek to prioritize cells that are aggregated in order to minimize boundary length.

Examples from literature

Moilanen, A., and Wintle, B. A. 2007. The boundary-quality penalty: a quantitative method for approximating species responses to fragmentation in reserve selection. *Conservation Biology*, 21: 355-364.

The theory and algorithm behind boundary length penalty is explained in section 2.4.1.

Input files

No special input files are needed in addition to the normal biodiversity feature layer grid sets and associated list file. The detailed combination of input files depends on the specific aims and components of your analysis.

Analysis stages and settings

Including BLP in your analysis is remarkably simple. All you need to do is assign a value for the penalty of increased boundary length. To do this, type the following line to your run settings file (section 3.3.2.3.)

```
BLP = [value]
```

The value of BLP should be a small decimal number. Try first a small value, e.g. 0.01, to perceive the effect of BLP to the solution. It is advisable to run the analysis with multiple BLP values to estimate the sensitivity of the solution to the value chosen.

When including BLP in the analysis, always use a warp factor of 1.

Strengths and weaknesses and further considerations

Note potential difficulties in interpretation of results if multiple aggregation methods are used simultaneously.

Link to tutorial

See tutorial exercise 4.

5.1.3 Distribution smoothing + Zonation

Planning problem to be solved

Accounting for connectivity of the protected area network from the perspective of species-specific dispersal ability. Distribution smoothing is a two-dimensional kernel smoothing, where the width of the smoothing kernel is determined by the estimated dispersal ability or scale of landscape use of the species in question. This option results a much more compact solution, where small isolated patches have been removed. Using distribution smoothing increases computation times marginally if at all.

Examples from literature

Moilanen, A., Franco, A. M. A., Early, R., Fox, R., Wintle, B., and Thomas, C.D. 2005. Prioritising multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society of London, Series B, Biological Sciences* 272: 1885-1891.

Moilanen, A. and Wintle, B.A. 2006. Uncertainty analysis favors selection of spatially aggregated reserve structures. *Biological Conservation,* 129: 427-434.

The theory and algorithm behind distribution smoothing is explained in section 2.4.2..

Pre-processing of inputs

To include distribution smoothing in your analysis, you need to determine the width of the dispersal kernel (a value for the distribution smoothing α) for each of your species, i.e. to what extent they can move in the landscape. Feed the α–values to the second column of your biodiversity feature list file.

Input files

- A set of biodiversity features grid layers (section 3.3.2.1.)
- A biodiversity features list file (section 3.3.2.2.) Enter the widths of dispersal kernels for each of your species in the second column of this file.
- A run settings file with appropriate settings (section 3.3.2.3.)

Distribution smoothing can be included in any analysis to induce aggregation into the protected area network. The detailed combination of input files depends on the specific aims and components of your analysis.

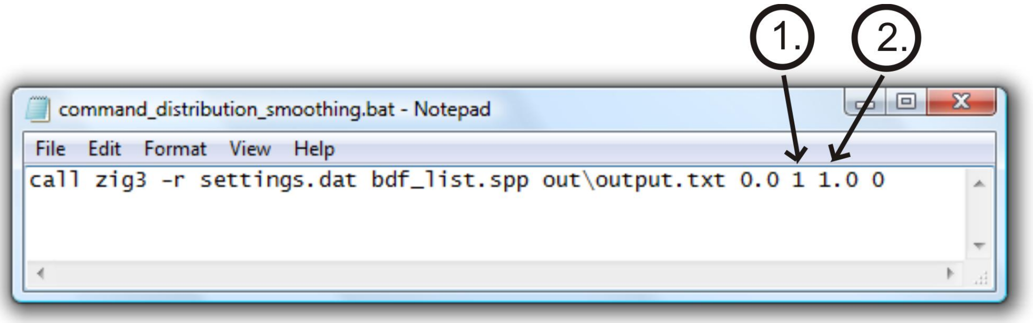
Analysis stages and settings

The adjustments to settings for distribution smoothing are not done in the run settings file, but straight to the command line.

1. Set the third last parameter of your command line call to 1 to indicate that distribution smoothing will be done.

2. Give a factor for multiplying the **species-specific α -values** as the second last

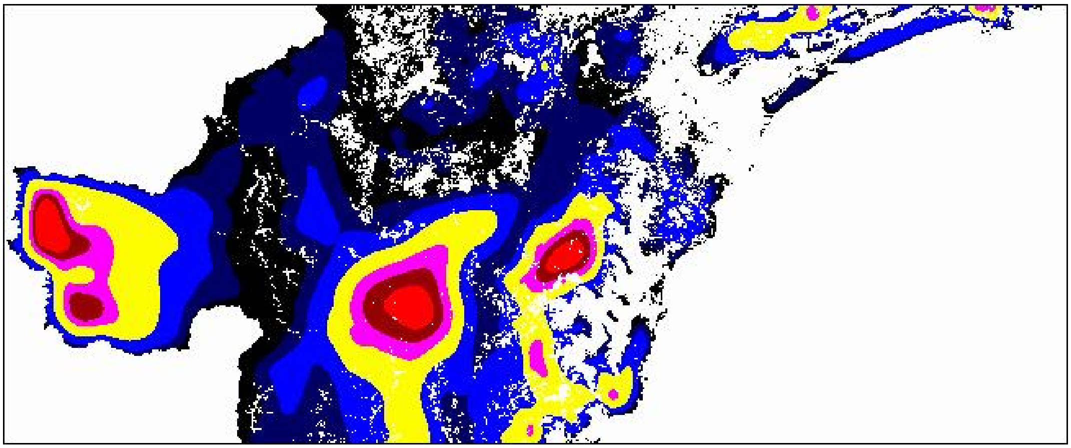
parameter in your call. α -values (species-specific widths of kernel) are in the second column of your biodiversity feature list file. The factor is useful if you are interested in running multiple solutions with, e.g. assuming several levels of dispersal capabilities, because the factor allows you to multiply all dispersal capabilities simultaneously. Thus you do not need to change the parameters manually in the species list file after each run. If you do not wish to multiply the α -values, set this factor to 1.



```
call zig3 -r settings.dat bdf_list.spp out\output.txt 0.0 1 1.0 0
```

Output and its interpretation

Note that using distribution smoothing should result a distinctively more aggregated solution compared to basic Zonation analysis. Thus these two solution should probably not have, e.g., a 99% overlap with each other. If your solution with distribution smoothing does not show clear aggregation, check the run settings for possible errors: for example, you could have α in different units than the cell size in the species distribution files.



Picture of a typical output map when distribution smoothing has been included in the analysis.

Strengths and weaknesses and further considerations

Strength: Distribution smoothing adds realism to the analysis, as it turns species-specific dispersal ability into a factor in the conservation planning process.

Multiple levels of connectivity, such as within and between populations of a single species, can be accounted for in a single Zonation run (see section 5.3.3.).

Link to tutorial

See tutorial exercise 4.

5.1.4 BQP Zonation

Planning problem to be solved

Accounting for species-specific effects of habitat fragmentation in conservation planning. Boundary-quality penalty (BQP) is a quantitative method that induces aggregation into reserve networks according to the needs of individual species.

Using the BQP decreases the biological quality of a land unit (grid cell) that is located close to the edge of the reserve, which results in a more highly aggregated optimal reserve structure.

Examples from literature

Moilanen, A., and Wintle, B. A. 2007. The boundary-quality penalty: a quantitative method for approximating species responses to fragmentation in reserve selection. *Conservation Biology*, 21: 355-364.

The theory and algorithm behind boundary quality penalty is explained in section 2.4.3.

Pre-processing of input files

To include BQP into your analysis, you first need to come up with a set of penalty curves that represent different types of responses to habitat fragmentation (see section 2.4.3.) and enter them into a BQP definition file (section 3.3.3.2.). Adjust your biodiversity feature list file by

- entering the row number of a correct BQP curve (in BQP definition file) for each biodiversity feature in the third column in order to link your features to BQP curves. Multiple features can be linked to the same penalty curve.
- indicating a suitable buffer size (in number of cells) for each species in the fourth column of the file.

Input files

To include BQP into your analysis, you need

- A BQP definition file (section 3.3.3.2.) containing responses to habitat fragmentation as penalty curves.
- A set of biodiversity features grid layers (section 3.3.2.1.)
- A biodiversity features list file (section 3.3.2.2.)
- A run settings file with appropriate settings (section 3.3.2.3.)

BQP can be included in any analysis to induce aggregation into the protected area network. The detailed combination of input files depends on the specific aims and components of your analysis.

Analysis stages and settings

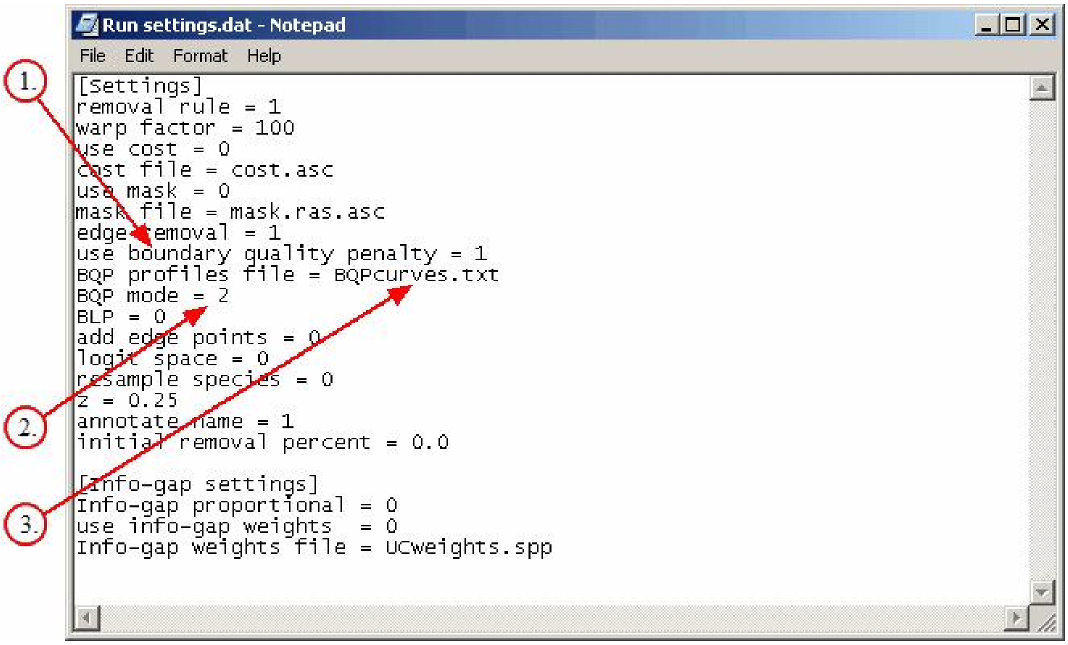
To run your analysis with BQP, type the following lines in your run settings file:

`use boundary quality penalty = 1`

`BQP profiles file = my_BQP_curves.txt`

`BQP mode = 1` if the data/no data matrix in all species distribution map files should be uniform and aligned, and that there are no differences between species in terms of which cells are considered potential habitat, and which are then used in BQP buffer calculations. In other words, all species would be dependent on the same general habitat type, such as forest OR

`= 2` if the data/no data matrixes are not uniform and aligned, and that the program needs to calculate species-specific buffers for each species/cell separately. Mode 2 is more realistic in the sense that fragmentation/loss in habitats that are not suitable for the species will not influence the value of the focal cell. But, mode 2 also requires longer computation times due to more complicated species-specific calculations.

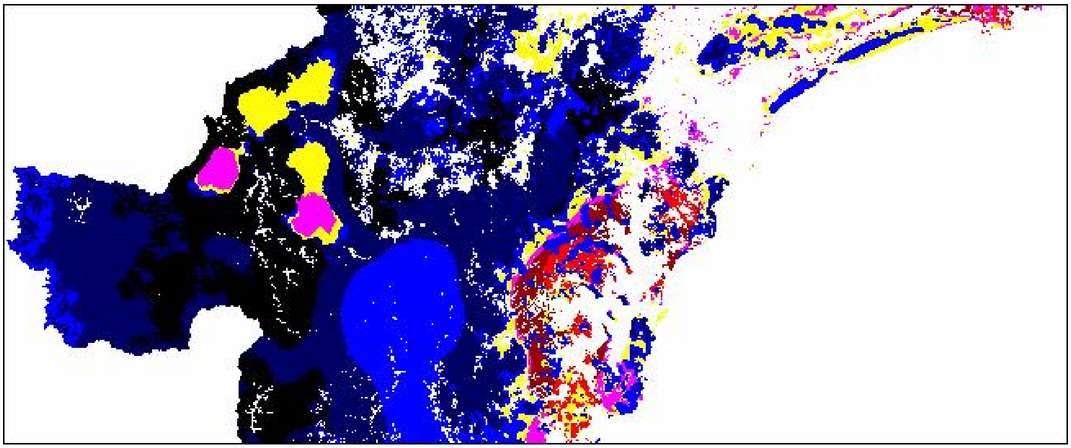


```
[Settings]
removal rule = 1
warp factor = 100
use cost = 0
cost file = cost.asc
use mask = 0
mask file = mask.ras.asc
edge removal = 1
use boundary quality penalty = 1
BQP profiles file = BQPcurves.txt
BQP mode = 2
BLP = 0
add edge points = 0
logit space = 0
resample species = 0
z = 0.25
annotate name = 1
initial removal percent = 0.0

[Info-gap settings]
Info-gap proportional = 0
use info-gap weights  = 0
Info-gap weights file = UCweights.spp
```

Output and its interpretation

Including BQP into your analyses most likely increases computation times significantly, especially if the species-specific buffer sizes are large (many cells). Thus it might be wise to reduce the data resolution if computation times start to increase in undesirable ways. It is also recommended to use a moderate to high warp factor for your preliminary runs and a low warp factor only for making the final solutions - low warp factors are recommended for final BQP runs. Using BQP should result in a distinctively more aggregated solution compared to basic Zonation analysis, at least in some part of the landscape. Thus, these two solutions should not have, e.g. 99% overlap with each other, unless your species are not influenced by habitat loss (as defined in the BQP file). If your solution with BQP shows signs of high fragmentation, check the run settings for possible errors. Note however, that a BQP solution may well include fragmented areas, if the data contains many species that are indifferent to fragmentation or even favor fragmented habitats.

Note that the use of BQP will produce species-specific curves that show lower proportions retained as compared to the basic (non-spatial) analysis. This is because use of the BQP implies that habitat loss and fragmentation will have negative consequences also for areas not yet lost. We emphasize that this does not mean that the solution developed by BQP is inferior to the basic analysis, but rather, the basic analysis gives an overly optimistic estimate of how much biodiversity value a fragmented landscape would retain.



Picture of typical output map when BQP has been included in the analysis.

Strengths and weaknesses and further considerations

Strength: BQP allows accounting for species-specific responses to habitat fragmentation, which makes it a rather realistic aggregation method.

Weakness: Individual considerations increase the need for biological data.

Please remember that even though there are no technical hindrance for using several aggregation methods in a single analysis, it may have unexpected effects and cause trouble for interpreting the results. Thus we recommend using only one aggregation method at a time.

Link to tutorial

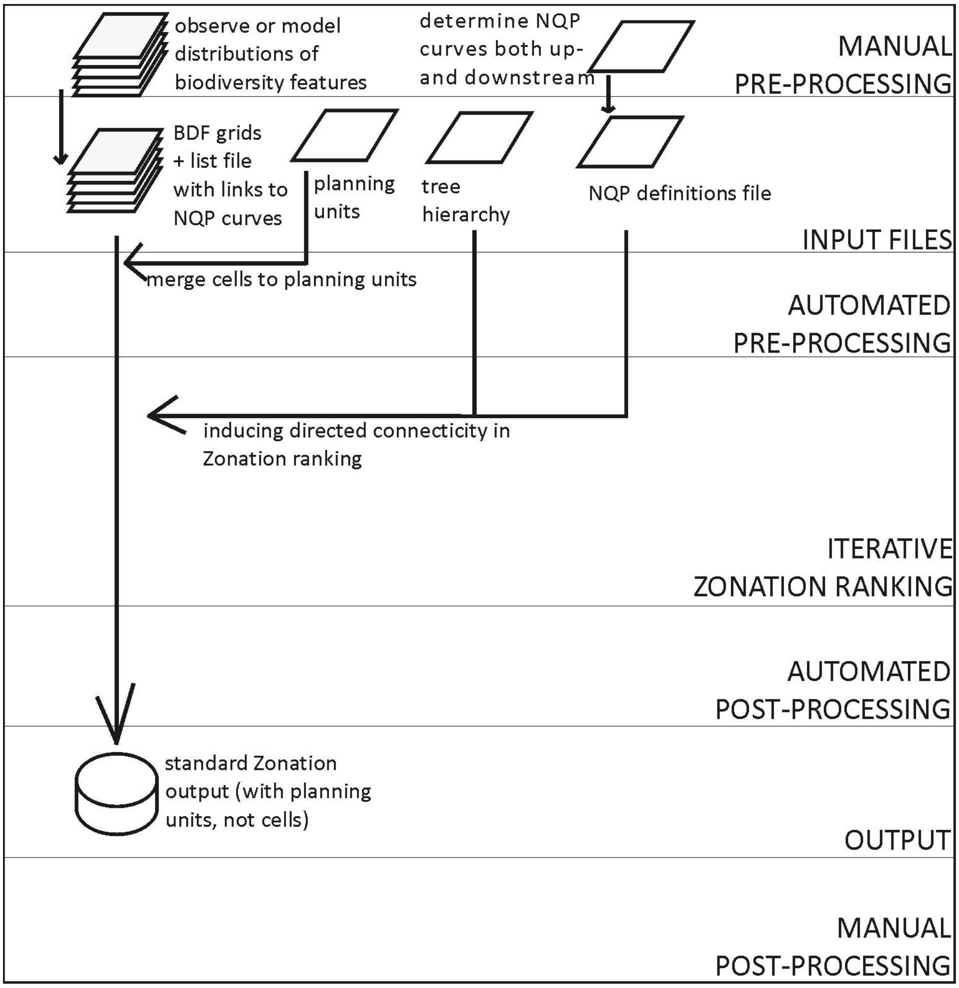
See tutorial exercise 4.

5.1.5 Directed freshwater connectivity

Planning problem to be solved

Identifying conservation priorities in freshwater systems where hydrological up- or downriver connectivity needs to be accounted for. Here, catchments are treated as planning units.

Process chart for the analysis


observe or model distributions of biodiversity features
determine NQP curves both up- and downstream
MANUAL PRE-PROCESSING
BDF grids + list file with links to NQP curves
planning units
tree hierarchy
NQP definitions file
INPUT FILES
merge cells to planning units
AUTOMATED PRE-PROCESSING
inducing directed connecticity in Zonation ranking
ITERATIVE ZONATION RANKING
AUTOMATED POST-PROCESSING
standard Zonation output (with planning units, not cells)
OUTPUT
MANUAL POST-PROCESSING

A process chart of an analysis with directed freshwater connectivity. Please note that only the compulsory and most commonly used optional analysis components are presented - you can combine different components according to your specific needs.

Examples from literature

Moilanen, A., Leathwick, J. and J. Elith. 2008. A method for freshwater conservation prioritization. *Freshwater Biology*, 53: 577-592.

Leathwick, J.R., Moilanen, A., Ferrier, S. and Julian, K. 2010. Community-level conservation prioritization and its application to riverine ecosystems. *Biological Conservation*, 143, 984-991.

Pre-processing of inputs

To induce directed freshwater connectivity into your prioritization, you first need to merge your data into planning units that consist of catchment areas. The hierarchy (in terms of direction of water flow) also needs to be identified. Finally, you should quantify species-specific responses to fragmentation of surrounding habitat and build sets of boundary quality penalty curves - one for responses to fragmentation upstream and one for downstream from the focal cell where the species occurs.

Input files

For an analysis with directed connectivity, you need

- A set of biodiversity feature grid layers (section 3.3.2.1.)
- A biodiversity feature list file (section 3.3.2.2.). Link all features to the correct penalty curves by entering the correct row number of the respective curve into the third (upstream) and fourth (downstream) columns.
- A planning unit layer, in which catchments are treated as planning units (section 3.3.3.11.)
- A tree hierarchy file describing the linkage between planning units (section 3.3.3.3.)
- A BQP definition file that describes both upstream and downstream connectivity responses. This means that each feature is assigned two penalty curves instead of one (section 3.3.3.2.).
- A run settings file with appropriate settings (section 3.3.2.3.)

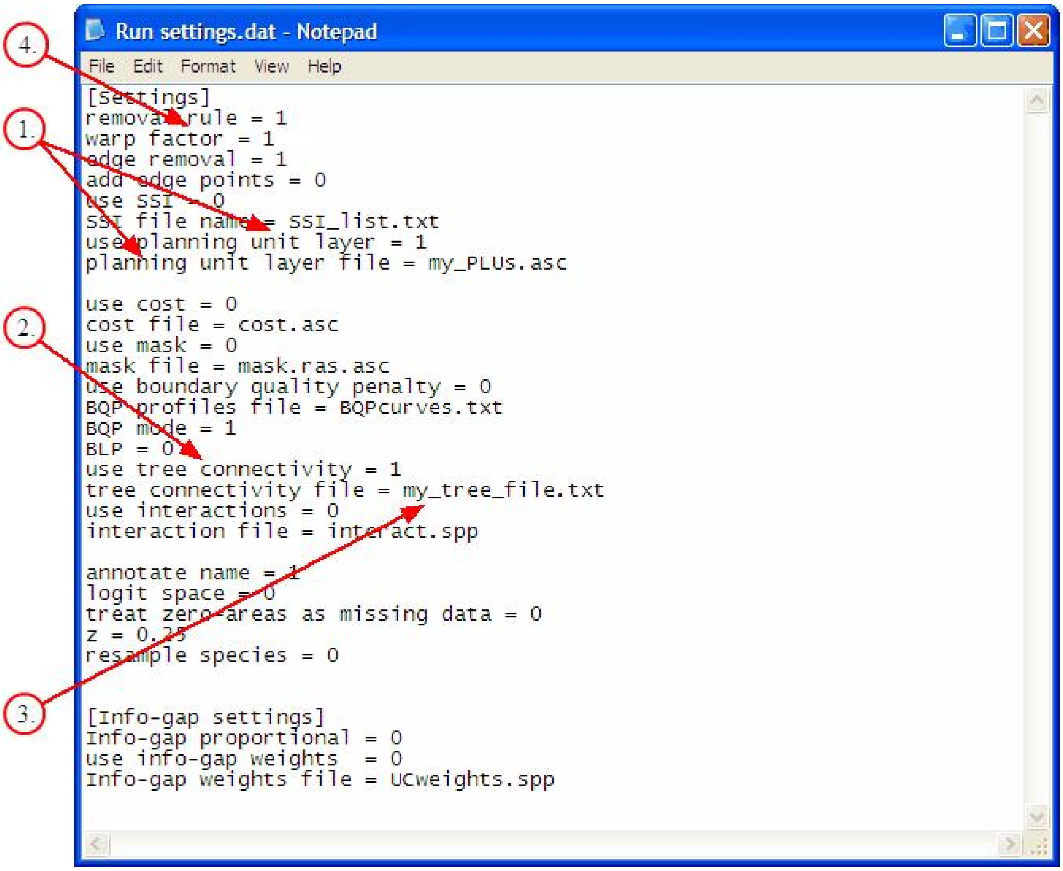
Depending on the specific aims and details of your analysis, you may want to include other input files, such as uncertainty layers, cost layer, landscape condition and retention layers etc.

Analysis stages and settings

To account for directed connectivity in your analysis, you need to adjust your run settings:

1. Set "`use planning unit layer`" to 1 in your run settings file and give the name of your planning unit layer file
2. Set "`use tree connectivity`" to 1 to indicate, that NQP will be used.
3. Also define the name of your NQP connectivity file in the run settings file.
4. Note that when planning units are used, the program will automatically set warp factor to

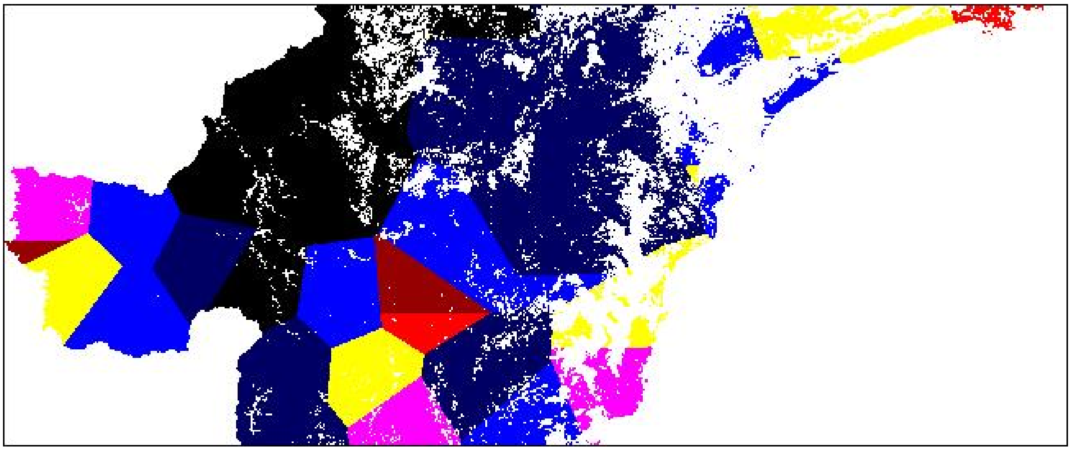
be one, regardless what has been defined in the run settings file.



```
[Settings]
removal rule = 1
warp factor = 1
edge removal = 1
add edge points = 0
use SSI = 0
SSI file name = SSI_list.txt
use planning unit layer = 1
planning unit layer file = my_PLUs.asc

use cost = 0
cost file = cost.asc
use mask = 0
mask file = mask.ras.asc
use boundary quality penalty = 0
BQP profiles file = BQPcurves.txt
BQP mode = 1
BLP = 0
use tree connectivity = 1
tree connectivity file = my_tree_file.txt
use interactions = 0
interaction file = interact.spp

annotate name = 1
logit space = 0
treat zero areas as missing data = 0
z = 0.25
resample species = 0


[Info-gap settings]
Info-gap proportional = 0
use info-gap weights  = 0
Info-gap weights file = UCweights.spp
```

Output and its interpretation

As NQP uses directed connectivity and planning units instead of singular grid cells, the output of this analysis should look significantly different in comparison to basic runs. The aggregation of areas depends on the average size of your planning units; large planning units lead to a high level of aggregation where as a small planning unit size gives more fine graded solutions.



Picture of output map when planning units and directed connectivity have been included in the analysis.

Note that the use of large planning units will automatically cause at least an apparent decrease in the quality of results. The reason for this is that large planning units will probably contain both areas that are good and bad for conservation. Consequently, the performance curves will suggest lower protection levels than what can be obtained if selection is based on individual grid cells. However, this effect can be counteracted by the relatively high levels of connectivity afforded by the use of the planning units. With respect to the computation time needed by Zonation, use of planning units is likely to cut computation times, with larger reductions produced by use of large planning units.

Strengths and weaknesses and further considerations

Strength: In the freshwater context, a relatively realistic analysis is possible with directed connectivity.

Weakness: Developing the files for freshwater connectivity is relatively complicated - the process has many steps where files can go wrong (planning units may end up being linked in a circle etc.). Zonation fixes some problems automatically and raises warnings. Please check the memo after these analyses for any warnings or error notes.

Link to tutorial

See tutorial exercise 8 for an example.

5.1.6 Matrix connectivity

Planning problem to be solved

Accounting for the extent to which multiple features influence each others connectivity. If your biodiversity features are species, you can use this analysis feature to include multiple ecological interactions. If you use community types as biodiversity features, you can account for the extent to which community types resemble each other, and facilitate dispersal or other processes that benefit from connectivity.

Examples from literature

Lehtomäki, J., Tomppo, E., Kuokkanen, P. Hanski, I., and A. Moilanen. 2009. Planning of forest conservation areas using high-resolution GIS data and software for spatial conservation prioritization. Forest Ecology and Management, 258: 2439-2449.

Pre-processing of inputs

To run an analysis that uses matrix connectivity to induce aggregation in the protected area network, you need to define the pair-wise interactions between your biodiversity features in terms of how the influence each others connectivity. You do not have to define connectivity interactions for all feature pairs, but can limit them to certain key features. See sections 2.4.5. and 3.3.3.4. and Lehtomäki et al. (2009) for details about the application of the matrix.

The features which you do include in the connectivity analysis should be listed first in your biodiversity features list file - in the same order as they appear in the matrix. For this analysis, the connectivity requirements are defined by feature-specific dispersal kernels that are used to transform occurrence levels to connectivity. Define widths of the dispersal kernels (dispersal α) of your biodiversity features and feed them in to the second column of your biodiversity feature list file.

Input files

To run Zonation with matrix connectivity, you need

- A set of biodiversity features grid layers (section 3.3.2.1.)
- A biodiversity features list file (section 3.3.2.2.)
- A connectivity similarity matrix (section 3.3.3.4.)
- A connectivity edge effect fix file (section 3.3.3.5.; this is optional)
- A run settings file with appropriate settings (section 3.3.2.3.)

Depending on the specific aims and details of your analysis, you may want to include other input files, such as uncertainty layers, cost layer, landscape condition and retention layers etc.

Analysis stages and settings

To run Zonation with matrix connectivity, you need the following lines in your run settings file:

```
[Community analysis settings]
load similarity matrix = 1
connectivity similarity matrix file = connectivity_matrix.txt
```
(the name of your connectivity matrix)
```
apply to connectivity = 1
```

If you wish to include edge effect fix, then you need an additional line:

```
connectivity edge effect fix file = fixfile_name.asc
```

If you wish to correct for habitat amounts within cells as well, use a cost layer with habitat proportions for each cell:

```
use cost = 1
cost file = habitatproportion_name.asc
```

Note that the parameter in the Zonation call multiplying the dispersal α-values modifies the distances used in matrix connectivity as well. See section 5.1.3. for more detailed instructions).

Strengths and weaknesses and further considerations

Developing the similarity matrices requires effort.

5.1.7 Cost efficiency analysis

Planning problem to be solved

Including cost into the selection of conservation priorities in order to maximize cost efficiency in terms of conservation value per cost unit. Cost can be integrated into any analysis by inputting a grid layer with cost information. Cost can refer to actual financial cost of acquiring the area for conservation, or it can represent the opportunity cost of not using the area for production of fish, timber or agricultural goods. In Zonation, conservation value in each grid cell is divided by cost of protecting that cell. This way, conservation value per cost unit is maximized in the solution.

Examples from literature

Leathwick, J. R., Moilanen, A., Francis, M., Elith, J., Taylor, P. Julian, K. and Hastie, T. 2008. Novel methods for the design and evaluation of marine protected areas in offshore waters.

For discussion about cost considerations in conservation planning, see
Arponen, A., Cabeza, M., Eklund, J., Kujala, H. and Lehtomäki, J. 2010. Costs of integrating economics and conservation planning. *Conservation Biology* 24: 1198–1204.

Pre-processing of inputs

It is possible to modify the effect to which cost influences the solution by altering the relative cost differences between cells. En example is provided by Leathwick et al. (2008). In their study, a cost layer was first created and then three modified cost layers were computed by introducing a parameter *a* as in

$$C_i^{\text{modified}} = \begin{cases} a, \\ a + 1 + \log_{10}(C_i), \end{cases}$$

with *a* assigned to cells in which conservation would imply no cost, and $a + 1 + \log_{10}(C_i)$ to the cells where a cost C_i would follow from conserving them. The larger the value of *a,* the smaller the differences between cells become, and thus cost has less influence in the solution. By using several values of *a,* you can also attain an idea about how sensitive your conservation solution is to the consideration of cost. In practice, this would mean computing several cost layers and then running the analysis a number of times, each of them with a different cost layer. The most convenient way to do this is to compile all the different analysis calls into a single batch file (see section 3.2.1.).

Input files

To include consideration of cost efficiency into your analysis, you need

- A set of biodiversity features grids (section 3.3.2.1.)
- A biodiversity feature list file (section 3.3.2.2.)
- A cost layer (section 3.3.3.6.).
- A run settings file with appropriate settings (section 3.3.2.3.)

Other input files depend on the details and other considerations of your analysis.

Analysis stages and settings

Consideration of cost can be integrated into any analysis. To do this, adjust your run settings by typing

```
use cost = 1

cost file = my_cost_layer.asc
```

into your run settings file.

An advisable strategy is to run the same analysis twice: first without and then with the cost layer.

Output and its interpretation

After running the analysis both with and without cost, it is possible to compare the outputs of the two runs and see how consideration of cost-efficiency affects the biodiversity value of the conservation solution. The thing to compare here are the representation curves (section 3.4.1.) from each analysis.

Strengths and weaknesses and further considerations

Strength: Taking cost into consideration is important in a real conservation situation. Targeting conservation action to areas with low opportunity cost can help alleviate conflicts of interest.

Weakness: If conservation priorities are chosen using only the cost-efficiency analysis, there is a risk that cost is actually driving the solution and priorities are strongly set based on low cost. This is the case especially if variation in cost is larger than variation of biodiversity value. See Arponen et al. 2010 for discussion on the topic.

Note: With Zonation, it is also possible to balance multiple land uses and related opportunity costs. This analysis is described in section 5.3.4.

Link to tutorial

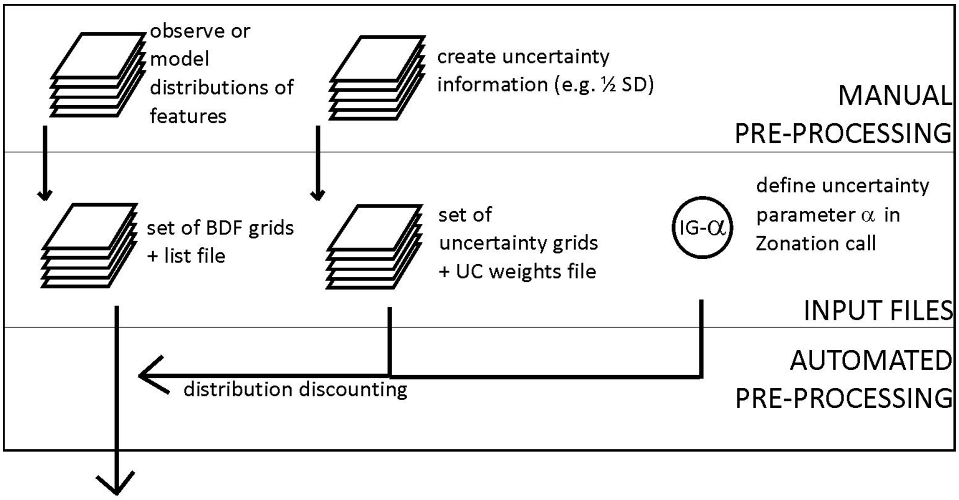
See tutorial exercise 6 for an example.

5.1.8 Analysis with uncertain inputs

Planning problem to be solved

Accounting for uncertainty in the occurrence data of biodiversity features. Uncertainty in the data can arise either from statistical distribution modeling or from threats to the continued persistence of the features. In both cases, uncertainty is related to whether the areas are actually suitable for the species or other biodiversity features, in a way that can guarantee their persistence in the long term. The application of information-gap theory in Zonation would give lower conservation values to areas where uncertainty is high. Consideration of uncertainty can be included in most analyses.

Process chart for the analysis


observe or model distributions of features
create uncertainty information (e.g. ½ SD)
MANUAL PRE-PROCESSING
set of BDF grids + list file
set of uncertainty grids + UC weights file
IG-α
define uncertainty parameter α in Zonation call
INPUT FILES
distribution discounting
AUTOMATED PRE-PROCESSING

Examples from literature

Moilanen, A., Runge, M. C., Elith, J., Tyre, A., Carmel, Y., Fegraus, E., Wintle, B., Burgman, M. and Ben-Haim, Y. 2006a. Planning for robust reserve networks using uncertainty analysis. *Ecological Modelling,* 199 (1): 115-124.

Moilanen, A., Wintle, B. A., Elith, J. and Burgman, M. 2006b. Uncertainty analysis for regional-scale reserve selection. *Conservation Biology,* 20: 1688-1697.

Moilanen, A. and Wintle, B.A. 2006. Uncertainty analysis favors selection of spatially aggregated reserve structures. *Biological Conservation,* 129: 427-434.

Carroll, C., Moilanen, A., and Dunk, J. 2010. Designing multi-species reserve networks for resilience to climate change: priority areas for spotted owl and localized endemics in the pacific North-West USA. *Global Change Biology*, 16: 891-904.

The theory and algorithms behind distributional uncertainty analysis are explained in section 2.5.

Pre-processing of inputs

A pre-processing step you often would take is fitting habitat suitability models to existing species data and creating spatial predictions of habitat suitability or occurrences of biodiversity features across the planning region.

To include distribution uncertainty in your analysis you first need to define the scale of uncertainty. As explained in section 2.5.1., this can be done by giving values to (1) the uncertainty parameter α and (2) to the species- and cell-specific relative error measure w_{sc} . Species-cell-specific relative errors are the ones given in the uncertainty map layers. **Using species-cell-specific errors (i.e. the uncertainty map layers) is optional, but you always have to give a value to the uncertainty parameter**.

The uncertainty parameter α determines the horizon of uncertainty in the data and is usually unknown. Thus you need to generate solutions with several α -values to determine how the spatial pattern behaves with increasing uncertainty. α can be either zero (= no uncertainty) or any positive value (distribution discounting) or any negative value (opportunity analysis).

If you are using α as the only measure of error (thus not using the distributional uncertainty map layers), it is important that the value of α is determined in relation to your data. For example, if your species data are probabilities of occurrence [0,1], the uncertainty parameter should be set to a reasonably small scale (e.g. $\alpha < 0.4$) to avoid complications, which may arise if all cells receive an effective discounted value of zero, leading to full loss of information from the distribution of the species.

To include species-cell-specific errors, you need to come up with

1. Distribution uncertainty map layers for each species. These layers show the relative magnitude of error (uncertainty) of species occurrence in each cell. Remember that the meaning of α must be interpreted with respect to the error measure you use. For example, if your error measure is the standard error of statistical prediction, then $\alpha=1$ essentially means subtracting one SD from the value of each cell.

2. Uncertainty analysis weights file containing a list of each distribution uncertainty layers and species-specific **error weights**. With error weights you can stress the data accuracy for certain, e.g. very rare, species. If no species uncertainty analysis weights are used, these parameters should be set to 1.0. Note that you can not use species-specific error weights without the distribution uncertainty map layers

Input files

To run a Zonation prioritization that accounts for uncertainty in occurrence of biodiversity features, you need

- A set of biodiversity feature grids (section 3.3.2.1.)
- A biodiversity feature list file (section 3.3.2.2.)
- A set of uncertainty map layers, one for each biodiversity feature (section 3.3.3.7.)
- An uncertainty map weights file (section 3.3.3.7.)
- A run settings file with appropriate settings (section 3.3.2.3.)

Depending on the specific aims and details of your analysis, you can include other input data as well.

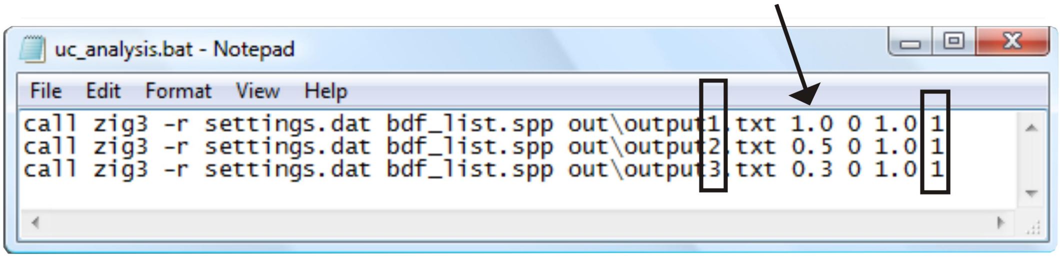
Analysis stages and settings

To set up an analysis that accounts for uncertainty, you need to adjust your run settings file. Type the following lines to the file:

`Info-gap proportional = 0` OR `= 1` depending on whether errors in species occurrences are uniform (value = 0) or proportional (value = 1). Uniform error is the default setting and works for most of the data sets, but in some cases it is more appropriate to use proportional errors (see e.g. Ben-Haim 2001).

```
use info-gap weights = 1

Info-gap weights file = my_IG_weights_file.txt
```

You also need to assign a value to the uncertainty parameter α in your program call (the fourth last parameter) in the command line. α -value determines the horizon of uncertainty in the data and is usually unknown. An advisable strategy is to generate solutions with several α -values to determine how the spatial pattern behaves with increasing uncertainty. α can be either zero (= no uncertainty) or any positive value (distribution discounting) or any negative value (opportunity analysis). One option to carry out the analyses is to write the calls for all of them in a single batch file. If you do this, remember to give different names for output files from different analyses. You should also set the last parameter of each command line to "1" to close the command prompt application and start over after each run.



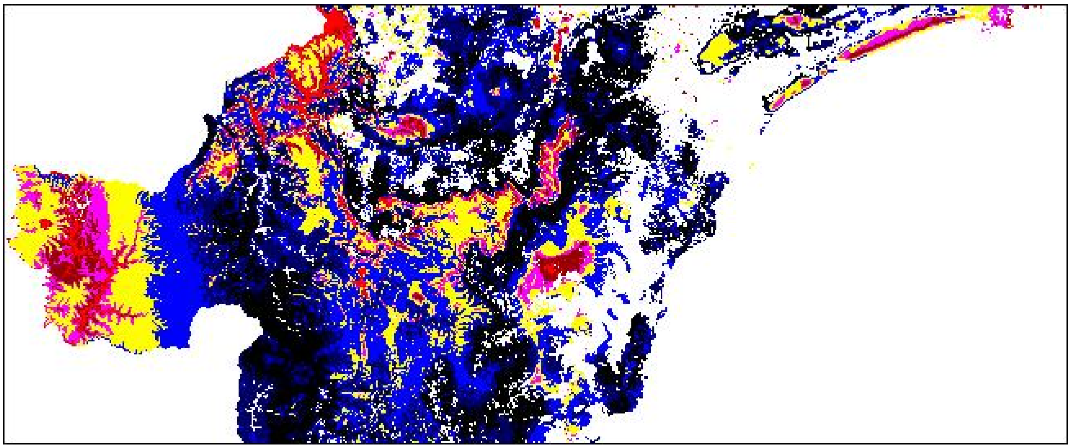
```
call zig3 -r settings.dat bdf_list.spp out\output1.txt 1.0 0 1.0 1
call zig3 -r settings.dat bdf_list.spp out\output2.txt 0.5 0 1.0 1
call zig3 -r settings.dat bdf_list.spp out\output3.txt 0.3 0 1.0 1
```

Other settings will depend on the specific aims and components of your analysis.

Sensitivity of the solution to the parameter settings can be analysed and alleviated by repeating the analysis several times with different values of α and checking for selection frequency of priority areas.

Output and its interpretation

The analysis produces a standard map of the landscape where reddish colors indicate sites that have both high species occurrence and high certainty. Depending on the amount of error in your data, the differences between the basic Zonation solution and distribution discounting can be significant - or not. In the Memo -window you can find more detailed information about the analysis. Note that as the program starts to run the analysis, it recalculates the species values based on the amount of uncertainty. Thus, for each species, the program first displays the absolute value in the whole landscape (sum over all cells) and then calculates how large fraction of this value can be expected to occur in the landscape with certainty. This value depends on the horizon of uncertainty parameter and on the level of uncertainty in the data.



Picture of our example landscape when uncertainty in species occurrences was included in the analysis.

Picture of the Memo -window when running distribution discounting.

Strengths and weaknesses and further considerations

Strength: The analysis explicitly acknowledges uncertainty and assigns high priorities to areas with low uncertainty about species occurrence or suitability.

Weakness: Parameter settings are subjective, which means that several analysis and consideration of solution sensitivity may be necessary.

Link to tutorial

See tutorial exercise 6 for an example of the analysis.

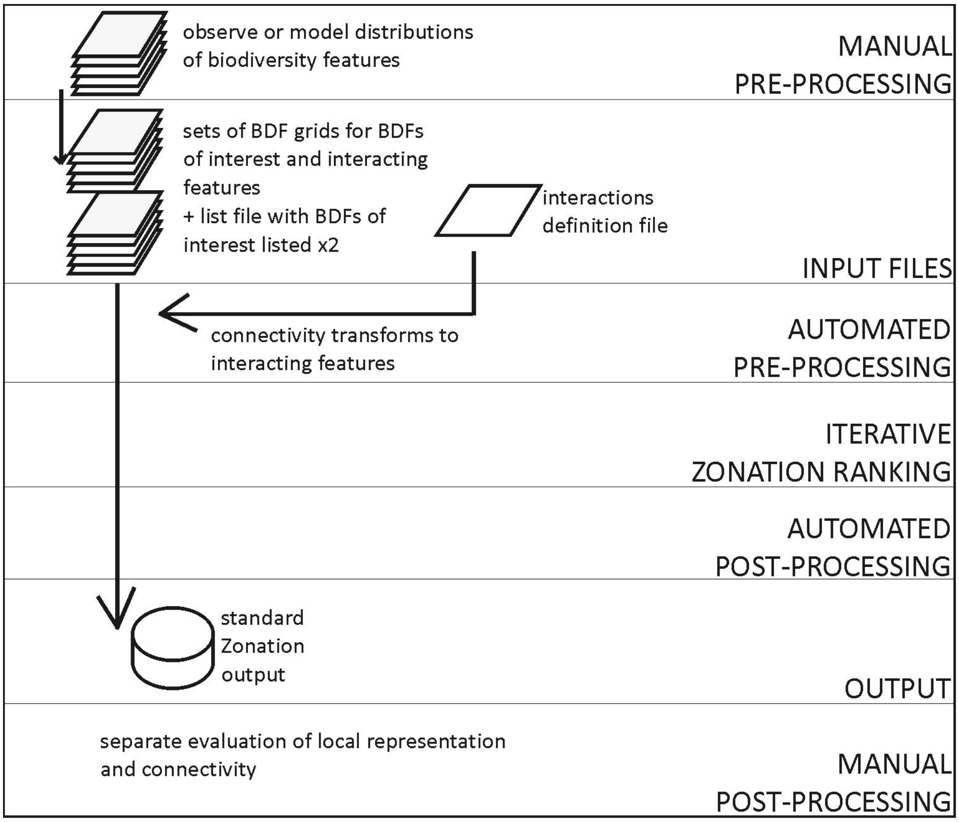
5.1.9 Ecological interactions

Planning problem to be solved

Incorporating ecological interactions into conservation prioritization. This method makes it possible to value cells not only by the presence of species within the cell, but also by its connectivity to important resources, such as a population of prey species, or to avoidable features, such as a population of a competitor species. This feature can also be used to induce connectivity of good quality habitats in the reserve network across different time steps to account for distribution shifts due to habitat management or climate change.

Zonation includes interactions by considering connectivity of suitable sites for the interacting features as a component of conservation value of each cell. For example, a cell in which an owl breeds can get a higher conservation value if there are cells with vole occurrences close by. The interactions can be either positive or negative.

Process chart


observe or model distributions of biodiversity features
MANUAL PRE-PROCESSING
sets of BDF grids for BDFs of interest and interacting features + list file with BDFs of interest listed x2
interactions definition file
INPUT FILES
connectivity transforms to interacting features
AUTOMATED PRE-PROCESSING
ITERATIVE ZONATION RANKING
AUTOMATED POST-PROCESSING
standard Zonation output
OUTPUT
separate evaluation of local representation and connectivity
MANUAL POST-PROCESSING

A process chart of an analysis with ecological interactions. Please note that only the compulsory and most commonly used optional analysis components are presented - you can combine different components according to your specific needs.

Examples from literature

Ecological interactions have been utilized to account for species interactions by

Rayfield, B., Moilanen, A. and Fortin, M.-J. 2009. Incorporating consumer-resource spatial interactions in reserve design. *Ecological Modelling*, 220: 725-733.

Interaction to enhance connectivity of good quality habitat between time steps have been utilized by

Carroll, C., Moilanen, A., and J. Dunk. 2010. Designing multi-species reserve networks for resilience to climate change: priority areas for spotted owl and localized endemics in the pacific North-West USA. *Global Change Biology*, 16: 891-904.

The theory and algorithm behind ecological interactions is explained in section 2.6.

Pre-processing of inputs

To run your analysis with ecological interactions included, determine connectivity requirements between interacting biodiversity features. They are defined in the interactions definition file by parameter *β.*

When compiling your biodiversity feature list file, please keep in mind the following points:

- **Every interaction changes the loaded, original distribution layer**. For example, lets assume you have two species (A and B), and you wish to run an analysis with the two original distributions plus the connection of species A distribution to species B distribution. To do this, you need to list species A twice in your species list file: The first layer will be used as it is (original distribution), and the second one will be transformed based on its connectivity to species B. If species A were listed only once, the landscape ranking would be done only based on species B distribution and species A connectivity to species B. Thus species A original distribution would not be included.
- **Make sure that you are not using already transformed layer to transform other layers**. Unless you absolutely want to. However, in that case the interpretation of results is outside the scope of this user manual.
- **Note that every file listed in the species list file will be used for landscape ranking**. If you want to use a layer to transform another layer, but not to be included into the analysis itself, you can do this by setting the weight of the layer to zero.

Input files

To consider ecological interactions in your analysis, you need

- A set of biodiversity features grid layers (section 3.3.2.1.). These should include layers for the species of interest and the interacting species, or observed present suitability and predicted future suitability of habitats for each species, depending on the analysis.

- A biodiversity feature list file (section 3.3.2.2.).

- An interactions definition file (section 3.3.3.8.) indicating which features interact and to what extent
- A run settings file with appropriate settings (section 3.3.2.3.)

Analysis stages and settings

To run Zonation with ecological interactions, adjust your run settings file by typing

```
use interactions = 1

interaction file = my_interactions.txt
```

Output and its interpretation

As usual, Zonation produces a map of cell ranking and all the visual and file output as with other analyses. In addition, the program prints to Memo all calculations related to using species interactions option. It is recommended to check from here that all interactions have been loaded correctly in the beginning of the analysis to avoid false results. With respect to interpretation of curves, assume a species (or whatever feature) has been entered into the same run as a raw distribution, a connectivity distribution and as transformed as a spatial interaction to a source. Then, the curve for the raw quality layer tells the fraction of local habitat quality retained, the curve for the smoothed layer tells about the fraction of connectivity retained, and the curve for the interaction tells about the fraction of the potential for spatial interaction retained.

Link to tutorial

No tutorial example is provided for this analysis.

5.2 Setups and interpretations for basic planning problems

In this section, we describe different ways to interpret the output of your analysis. The interpretations can be applied to any analysis setup.

5.2.1 Selecting conservation areas

Planning problem to be solved

Identifying the most valuable areas of the landscape to form a conservation area network. An optimal network displays a balance of high quality areas and connectivity for conservation features. Maximizing representation of biodiversity features is a central goal in systematic conservation planning. With this approach, it is also possible to select areas for other forms of conservation action than assigning protected areas. This analysis can be adjusted to account for connectivity in various ways. It can also account for cost, uncertainty, landscape condition and retention, suitability for alternative land uses and other factors that are present in real life conservation planning.

Examples from literature

Moilanen, A., Franco, A. M. A., Early, R., Fox, R., Wintle, B., and Thomas, C.D. 2005. Prioritising multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society of London, Series B, Biological Sciences* 272: 1885-1891.

Leathwick, J. R., Moilanen, A., Francis, M., Elith, J., Taylor, P. Julian, K. and Hastie, T. 2008. Novel methods for the design and evaluation of marine protected areas in offshore waters. *Conservation Letters*, 1: 91-102.

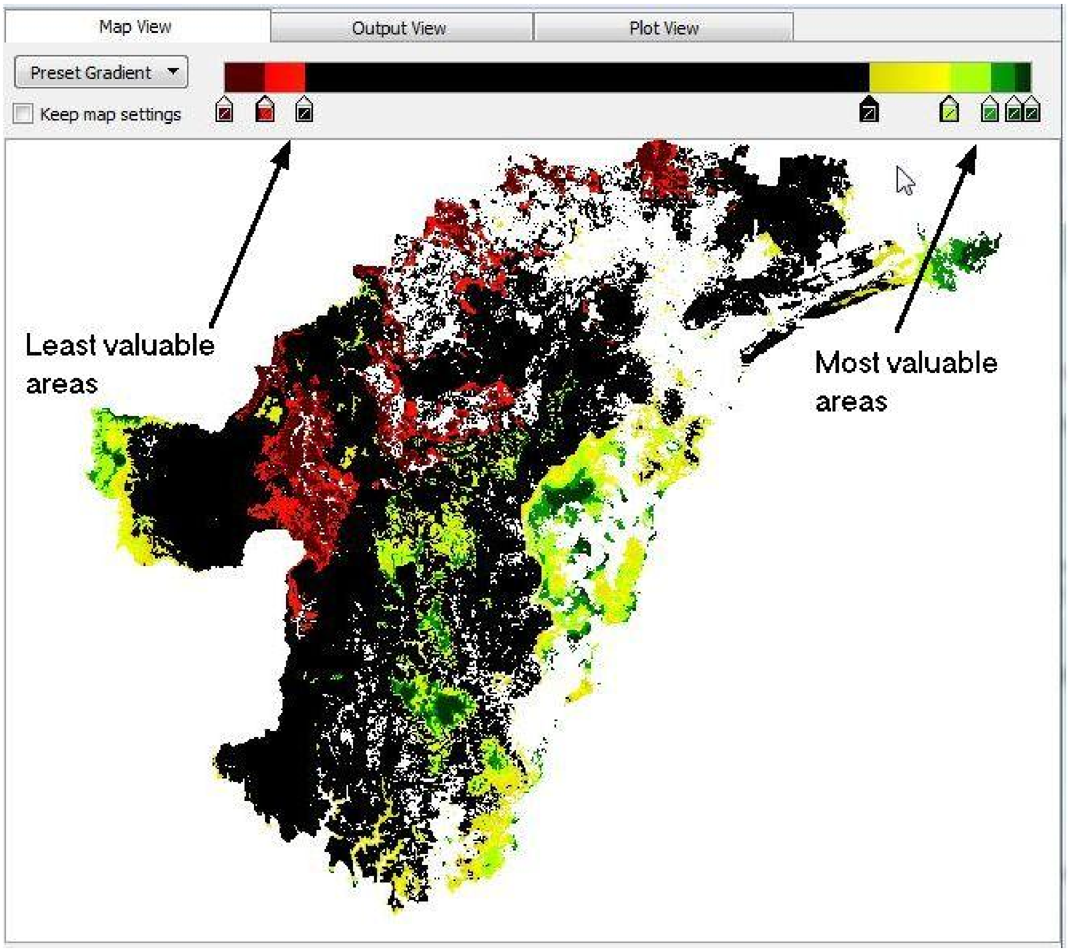
Analysis stages and settings

The analysis to identify most valuable areas for conservation can be performed with various types of input data, as well as variable setup combinations and degrees of complexity. See simple Zonation (section 5.1.1.) for necessary input files and settings for a simple non-spatial analysis. Input files, pre-processing, additional features used and interpretation of output depend on the specific goals and data in each case. Analysis setups to include spatial and ecological factors are explained in the later sections.

Output and its interpretation

A central stage in an analysis to select conservation areas is to look for the most valuable areas in the top fraction of Zonation priority rank map. Zonation produces representation curves for biodiversity features' representation levels as cells are removed. These curves can be used to verify the quality of the top priority areas for conservation.

Zonation GUI allows adjusting the colour gradient on the output map to match the specific planning situation. It is possible to show the most meaningful proportions of the most valuable areas as well as the least valuable areas on a single map (see figure below for an example and section 4.3.1. for advice).


Map View
Output View
Plot View
Preset Gradient
☐ Keep map settings
Least valuable areas
Most valuable areas

Strengths and weaknesses and further considerations

Strengths: Using this approach can provide a valuable reference baseline for replacement cost analysis.

Weaknesses: In most cases, the conservation area network cannot be created from a clean sheet. Selecting the most valuable areas across the entire landscape may therefore not provide the most appropriate solution, when previous limitations, land use plans and existing protected areas are not accounted for.

Link to tutorial

See exercise 1 for a tutorial example of the basic use of Zonation. Exercise 2 provides an example of species weighting.

5.2.2 Identifying least valuable areas for conservation

Planning problem to be solved

Identifying areas with low conservation value and therefore most suitable for other land uses. This approach can be applied when land needs to be assigned for purposes that do not support biodiversity. By identifying the least valuable areas, it is possible to guide land use so that biodiversity is adversely affected as little as possible.

Examples from literature

Gordon, A., Simondson, D., White, M., Moilanen, A., and Bekessy, S.A. 2009. Integrating conservation planning and land-use planning in urban landscapes. *Landscape and Urban Planning*, 91: 183-194.

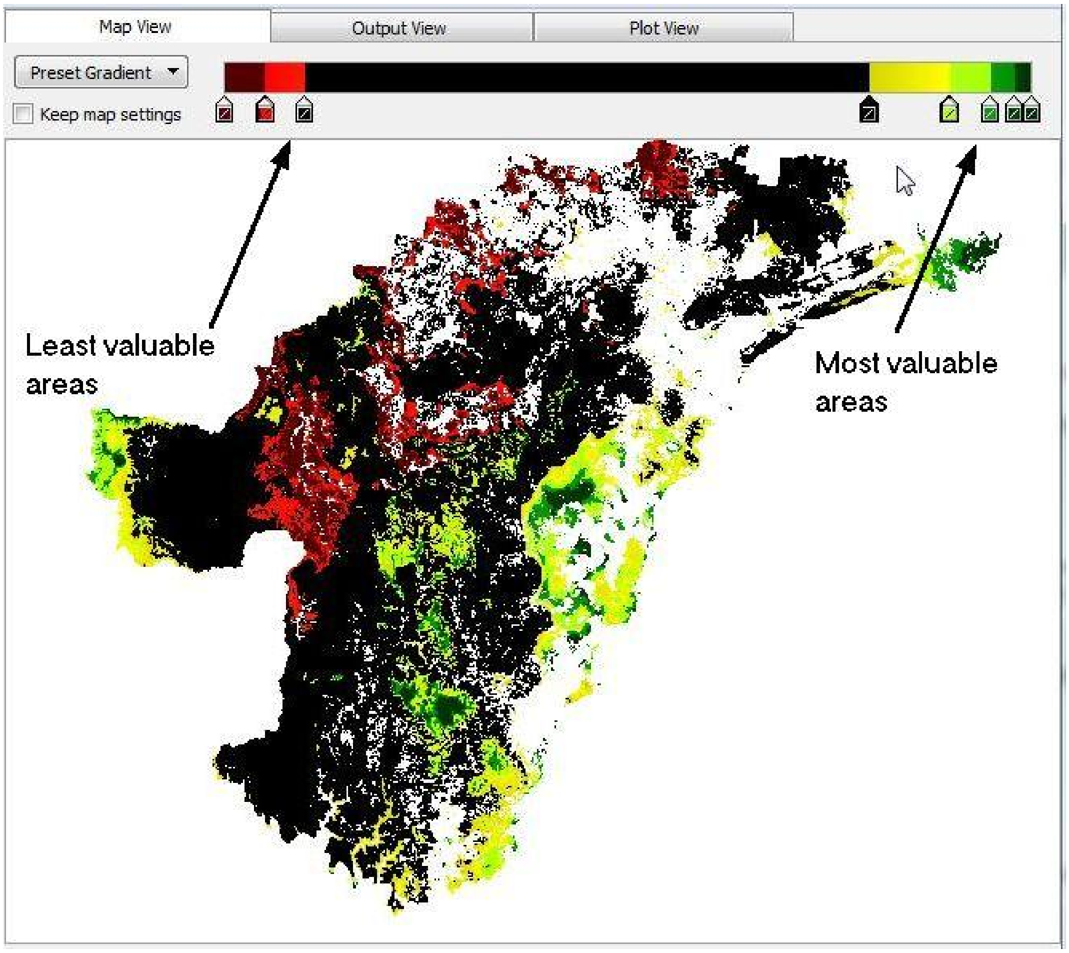
Analysis stages and settings

Identifying least valuable areas is, in fact, performed in exactly similar way as selecting conservation areas. The only difference is in the interpretation of the output: here, one should look for the lowest fraction of priority ranking. Lowest priorities are assigned to sites that contribute least conservation value to a network. This interpretation can be applied to any Zonation analysis, simple or complex. See simple non-spatial Zonation (section 5.1.1.) for necessary input files and settings to run a non-spatial analysis. You can add spatial and ecological considerations according to your needs. Setups for more complex analyses are described in the following sections.

Output and its interpretation

When identifying areas not that important for biodiversity, it is important to bear in mind that the results reflect only those aspects of biodiversity that were included in the analysis.

Zonation GUI allows adjusting the colour gradient on the output map to match the specific planning situation. It is possible to show the most meaningful proportions of the most valuable areas as well as the least valuable areas on a single map (see figure below for an example and section 4.3.1. for advice).


Map View
Output View
Plot View
Preset Gradient
☐ Keep map settings
Least valuable areas
Most valuable areas

Strengths and weaknesses and further considerations

An alternative approach to identify areas for other land uses is to optimize multiple land use purposes simultaneously (see section 5.3.4.).

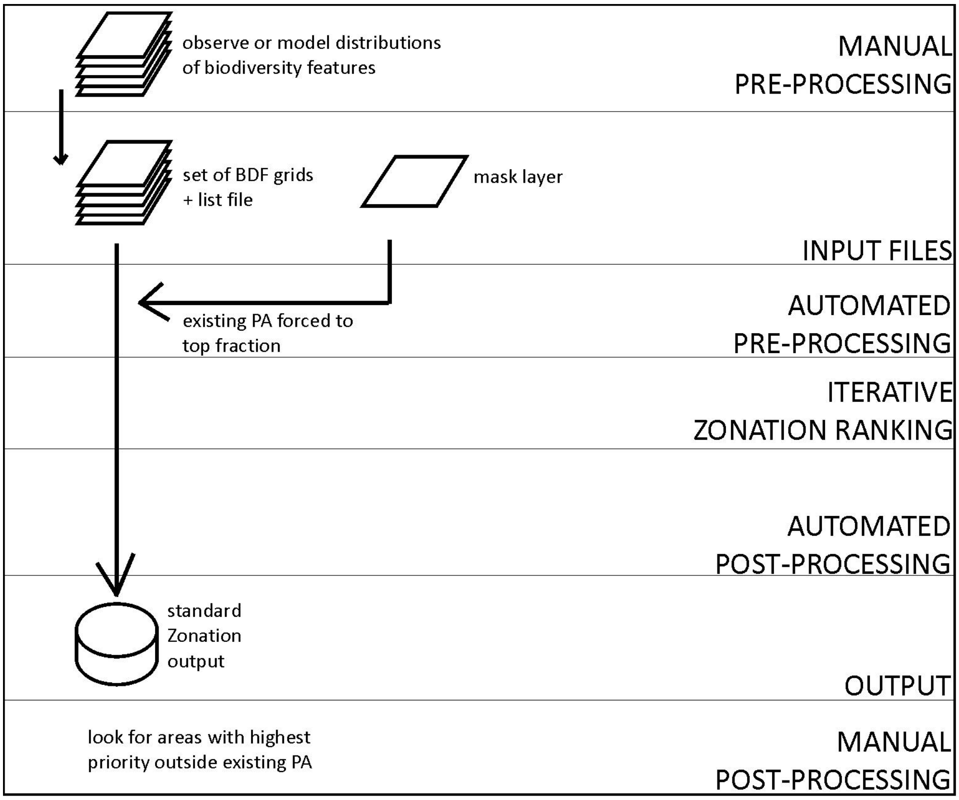
5.2.3 Expanding conservation areas

Planning problem to be solved

Identifying areas that would best complement the already existing conservation area network, i.e. they would add most conservation value. The main focus is on the quality of the new areas and the overall connectivity of the conservation area network.

This variant of selecting conservation areas can be applied in any of the more complicated analyses. The key component in the variant is the area inclusion mask that forces the existing protected areas into the top fraction of the solution.

Process chart for the analysis


observe or model distributions of biodiversity features
MANUAL PRE-PROCESSING
set of BDF grids + list file
mask layer
INPUT FILES
existing PA forced to top fraction
AUTOMATED PRE-PROCESSING
ITERATIVE ZONATION RANKING
AUTOMATED POST-PROCESSING
standard Zonation output
OUTPUT
look for areas with highest priority outside existing PA
MANUAL POST-PROCESSING

A process chart of the analysis for expanding a conservation area network. Please note that only the compulsory analysis components are presented - you can combine different components according to your specific needs.

Examples from literature

Lehtomäki, J., Tomppo, E., Kuokkanen, P. Hanski, I., and A. Moilanen. 2009. Planning of forest conservation areas using high-resolution GIS data and software for spatial conservation prioritization. Forest Ecology and Management, 258: 2439-2449.

Kremen, C., Cameron, A., Moilanen, A., Phillips, S., Thomas, C.D., Beentje, H., Dransfeld, J., Fisher, B.L., Glaw, F., Good, T., Harper, G., Hijmans, R.J., Lees, D.C., Louis Jr., E., Nussbaum, R.A., Razafimpahanana, A., Raxworthy, C., Schatz, G., Vences, M., Vieites, D.R., Wright, P.C. and Zjhra, M.L. 2008. Aligning conservation priorities across taxa in Madagascar, a biodiversity hotspot, with high-resolution planning tools. *Science*, 320: 222-226.

Input files

Compulsory files to perform this analysis include

- A set of biodiversity feature grids (section 3.3.2.1.)
- Associated biodiversity feature list file (section 3.3.2.2.)
- A run settings file with appropriate settings (section 3.3.2.3.)
- Removal mask layer to mask in existing conservation areas (section 3.3.3.9.)

Optional input files, depending on the specific setting and considerations of each case, can include

- Cost layer (section 3.3.3.6.)
- Uncertainty layers and weights file (section 3.3.3.7.)
- Set of landscape condition layers (section 3.3.3.14.)
- Set of landscape retention layers (section 3.3.3.15.)
- Matrix of community similarity, if your biodiversity features are communities (section 3.3.3.4.)

In addition to the inputs above, one would probably wish to define connectivity transforms of responses using some of the feature-specific connectivity methods available in Zonation (sections from 5.1.3. to 5.1.6.)

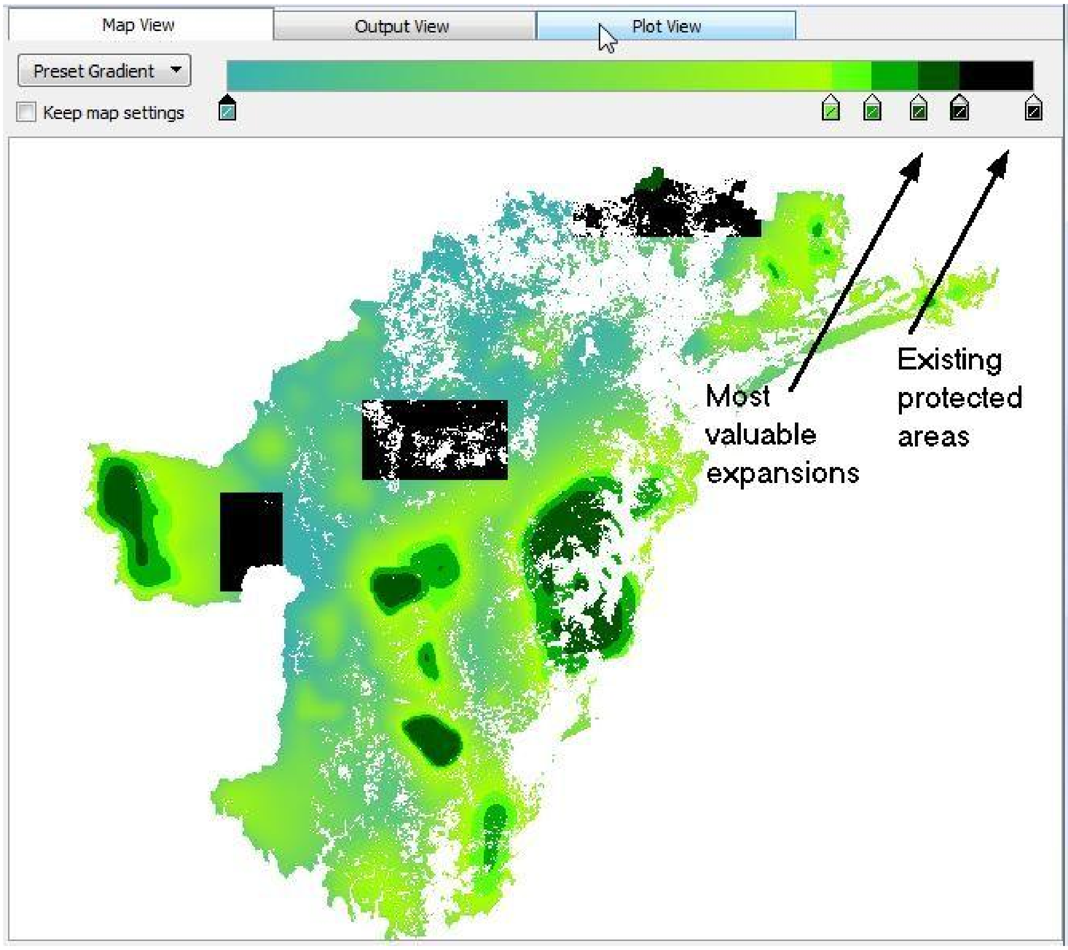
Analysis stages and settings

Depends on the specific aims of your analysis. However, typically at least connectivity responses would be included to model connectivity to existing protected areas.

Output and its interpretation

Existing protected areas were forced into the top fraction of the priority ranking with area inclusion mask. To identify areas to complement the existing protected area network most efficiently, look for areas with the highest ranking outside existing protected areas.

Zonation v.3.0 GUI allows adjusting the colour gradient so that you can easily visualise the location of best areas outside the areas that are already protected (see section 4.3.1. for advice).


Map View
Output View
Plot View
Preset Gradient
☐ Keep map settings
Most valuable expansions
Existing protected areas

Strengths and weaknesses and further considerations

Strength: provides a clear answer to a clear question.

Weakness: Setup may turn out relatively complicated.

Link to tutorial

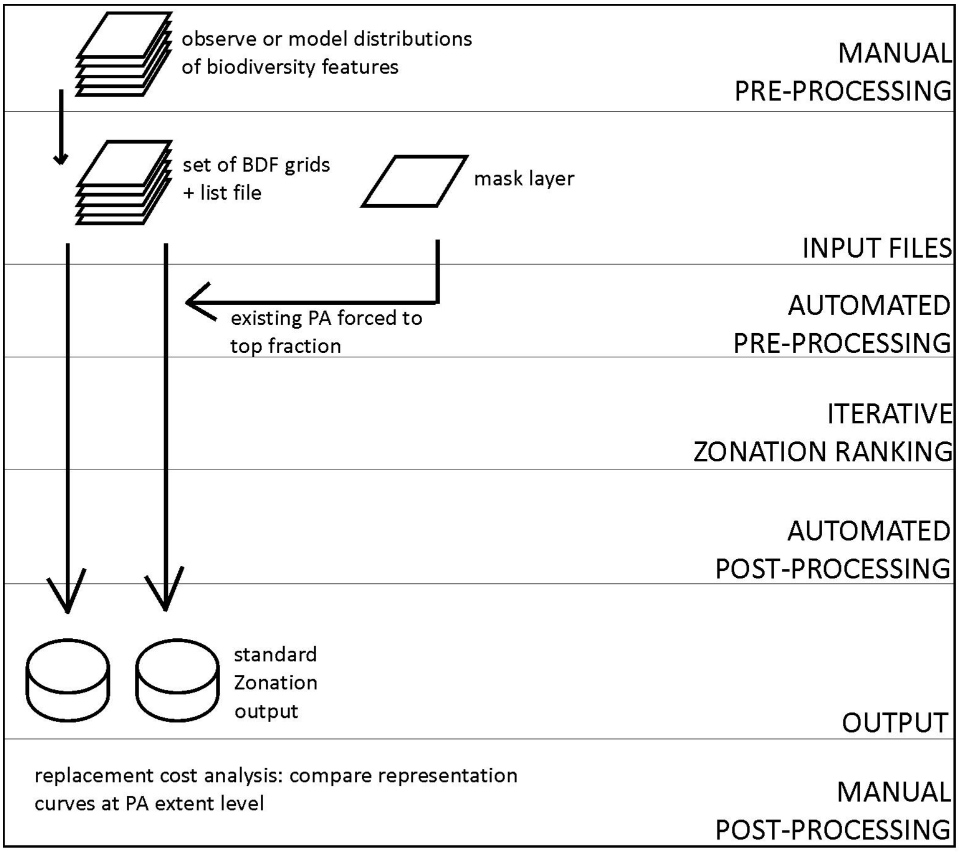
See Exercise 7 for an example.

5.2.4 Evaluating existing/proposed conservation areas

Planning problem to be solved

Evaluating the conservation value of existing or proposed networks of protected areas in terms of representation and/or connectivity. An essential component of this analysis is replacement cost (see section 2.7.) - how much of the maximum possible conservation value is lost if a given network is protected. The question can be turned around, as in how much additional funds or land is needed to achieve the same conservation value as in the unconstrained top fraction.

Process chart for the analysis


observe or model distributions of biodiversity features
MANUAL PRE-PROCESSING
set of BDF grids + list file
mask layer
INPUT FILES
existing PA forced to top fraction
AUTOMATED PRE-PROCESSING
ITERATIVE ZONATION RANKING
AUTOMATED POST-PROCESSING
standard Zonation output
OUTPUT
replacement cost analysis: compare representation curves at PA extent level
MANUAL POST-PROCESSING

A process chart of the analysis for evaluating a conservation area network. Please note that only the compulsory analysis components are presented - you can combine different components according to your specific needs.

Examples from literature

Cabeza, M. and Moilanen, A. 2006. Replacement cost: a useful measure of site value for conservation planning. *Biological Conservation,* 132: 336-342.

Leathwick, J. R., Moilanen, A., Francis, M., Elith, J., Taylor, P. Julian, K. and T. Hastie. 2008. Novel methods for the design and evaluation of marine protected areas in offshore waters. *Conservation Letters*, 1: 91-102.

Kremen, C., Cameron, A., Moilanen, A., Phillips, S., Thomas, C.D., Beentje, H., Dransfeld, J., Fisher, B.L., Glaw, F., Good, T., Harper, G., Hijmans, R.J., Lees, D.C., Louis Jr., E., Nussbaum, R.A., Razafimpahanana, A., Raxworthy, C., Schatz, G., Vences, M., Vieites, D.R., Wright, P.C. and Zjhra, M.L. 2008. Aligning conservation priorities across taxa in Madagascar, a biodiversity hotspot, with high-resolution planning tools. *Science*, 320: 222-226.

Input files

Necessary input files for the analysis include

- A set of biodiversity feature grids (section 3.3.2.1.)
- A feature list file (section 3.3.2.2.)
- A run settings file with appropriate settings (section 3.3.2.3.)
- A removal mask layer to force existing or proposed protected areas (section 3.3.3.9.)

Analysis stages and settings

To do a replacement cost analysis, you actually need to compute a priority ranking on the same biodiversity feature dataset twice.

First, run a normal Zonation prioritization without a mask file, as when selecting conservation areas. In all other respects, your settings should be the same as for the second run. The output from the first run is the unconstrained solution that serves as a reference for replacement cost analysis.

Second, run Zonation with otherwise the same settings as the first time, but now including the area inclusion mask file with areas in the current or proposed network forced in to the top fraction. For this, you need to adjust your run settings:

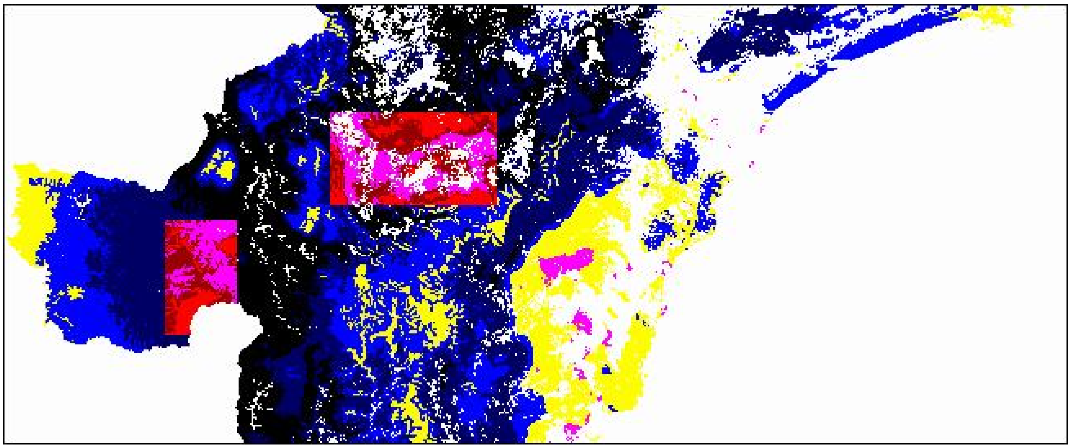
```
use mask = 1

mask file = my_current_network.asc
```

The detailed settings of your analysis depend on the combination of analysis components you want to include.

Output and its interpretation

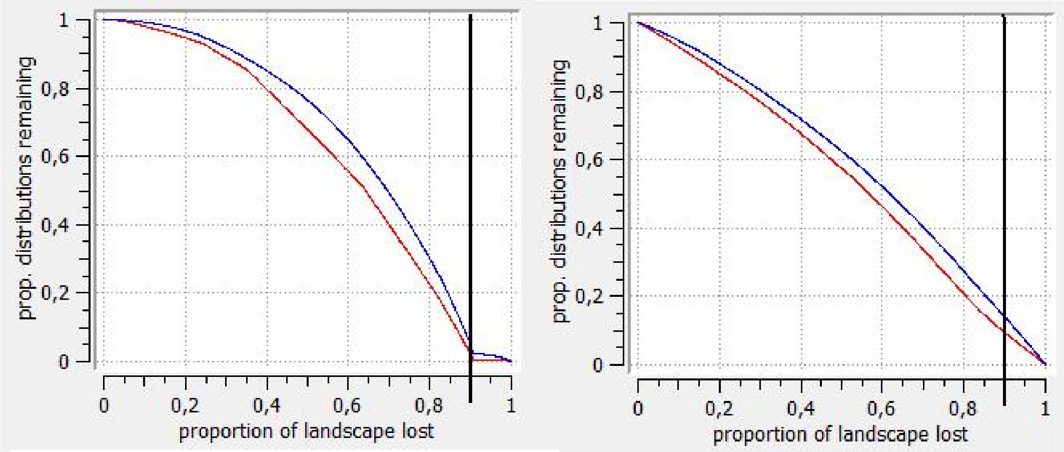
To evaluate the difference between constrained and unconstrained solution in their conservation value, go to the representation curves of both runs (section 3.4.1.). Compare the curves at the level of cell removal that corresponds to the area of your existing or proposed protected area network.

The effects of running the analysis with a mask file should be clearly seen from your output map. The included areas should receive the highest values in landscape ranking where as the excluded areas receive the lowest values.



Picture of our example landscape when two large squarish areas have been forcibly included to the final solution with a mask file.

Also the distribution curves in **Plots tab** may show some changes depending on the areas that have been included/excluded. Note that in many cases the use of mask files results a suboptimal solution. In other words, the program cannot select the best possible solution, because it is forced to either exclude biologically valuable areas from or include poor quality areas into the top fraction. The picture below shows an example of comparing the performance curves between a solution where certain areas are forced into the top fraction (left) and an unrestricted solution (right). The difference between the proportion of distributions remaining at a given fraction of landscape can be considered the biological inclusion cost.



Picture of distribution curves when the program has been forced to include low quality areas to the solution above. Note the changes at the end of the curves, clearly demonstrating that the forcibly included areas were not what one would have ideally chosen.

By comparing the .curves.txt -files that are produced after each run, it is possible to evaluate the costs induced by the usage of inclusion/exclusion masks (see also section 2.7.).

5.2.5 Targeting of incentive funding

Planning problem to be solved

Obtaining information about (i) how to target conservation marketing or (ii) which site to choose amongst multiple sites offered for conservation or other similar decisions. Here, as the decision maker is not in search of the optimal solution considering the whole landscape, Zonation ranking outputs are interpreted in a different manner than usually.

Analysis stages and settings

This approach can be utilized with any setup combination. The components to use depend on the specific aims and details of your analysis.

Output and its interpretation

The relevant output for this approach is the weighted range size corrected richness (wrscr) raster (see section 3.4.1.), which is one of the maps Zonation automatically outputs from each analysis. It illustrates intrinsic conservation values of sites, irrespective of complementarity or connectivity in relation to other sites. Choices between a certain number of sites is possible by comparing their wrscr values. Range sizes and weighting of species or other biodiversity features and the combination of analysis components will affect the wrscr scoring.

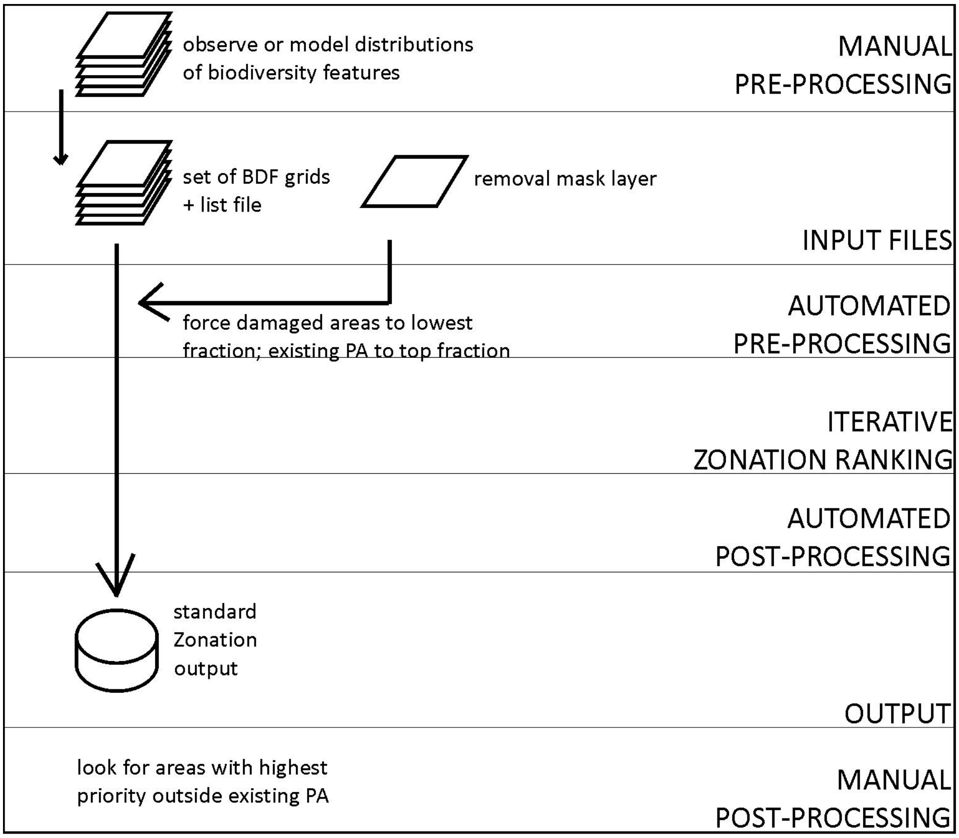
As the wrscr scoring does not account for the relative values of the site in the landscape or network context, it is advisable to examine also the basic priority rank map and base decisions on both maps.

5.2.6 Offsetting and targeting of compensation measures

Planning problem to be solved

Identifying targets to compensate (offset) a biodiversity loss (due to economic activity) in one location by conservation action at another location.

Process chart for the analysis


observe or model distributions of biodiversity features
MANUAL PRE-PROCESSING
set of BDF grids + list file
removal mask layer
INPUT FILES
force damaged areas to lowest fraction; existing PA to top fraction
AUTOMATED PRE-PROCESSING
ITERATIVE ZONATION RANKING
AUTOMATED POST-PROCESSING
standard Zonation output
OUTPUT
look for areas with highest priority outside existing PA
MANUAL POST-PROCESSING

A process chart of the analysis for targeting compensatory measures. Please note that only the compulsory analysis components are presented - you can combine different components according to your specific needs.

Examples from literature

This analysis has not been published. For discussion on aspects to consider when planning offsetting, see

Moilanen, A., van Teeffelen, A.J.A., Ben-Haim, Y. & Ferrier, S. 2009. How much compensation is enough? A framework for incorporating uncertainty and time discounting when calculating offset ratios for impacted habitat. *Restoration Ecology* 17: 470-478.

Input files

To identify areas to target compensation measures after habitat conversion due to economic activity, you need

- A set of biodiversity features grids (section 3.3.2.1.)
- A removal mask layer (section 3.3.3.9.). Assign lowest mask values (e.g. 1) to cells that are adversely impacted by the planned activity to exclude them from the solution. Highest mask values (e.g. 3) should be assigned to cells that are already protected, so that they will be included in the top rank. The rest of the landscape should be assigned intermediate mask values (e.g. 2).
- A run settings file with appropriate settings (section 3.3.2.3.)

Depending on the specific aims and details of your analysis, you may want to include considerations of cost, alternative land uses, landscape condition and retention etc.

Analysis stages and settings

The basic setup for this analysis is rather simple. The only adjustments you need to make to your run settings file are

`run mask = 1`

`mask file = my_mask_file.asc` (name of your removal mask layer file)

The detailed setup will depend on the combination of components you want to include.

Output and its interpretation

Look for sites with top ranking outside areas that are already protected (and masked in the top fraction). Those would be the best sites for targeting compensation measures in terms of how they complement the existing protected area network.

Strengths and weaknesses and further considerations

This analysis corresponds to the weak sustainability assumption (see Moilanen et al. 2008), which does not require that compensation is near the damaged area nor targeted to the exactly same habitats. If strong sustainability is required, then analysis needs to be limited to damaged habitats and connectivity to damaged areas included (this is liable to be inefficient use of resources).

Link to tutorial

We do not provide a tutorial example for offsetting. Please refer to exercise 7 for a tutorial example for using mask files.

5.3 Full analysis setups

In this section, we present examples of full analysis setups in different types of planning contexts. The examples presented here can further be combined - but proceeding with care is advisable. Section 5.3.9. provides some guidelines for creative combinations of analysis components.

5.3.1 Community level analysis

Planning problem to be solved

In the examples so far, as often in conservation planning, biodiversity data has been a set of individual species. The community level analysis operates on habitat types or community classes. The aim in the community level priority setting is to maximize complementary representation of habitat or community types, weighted by their richness. The complementarity approach utilizes information about similarity, i.e. shared species between community classes.

Examples from literature

The theoretical basics of community level conservation prioritization is explained by

Arponen, A., Moilanen, A., and S. Ferrier. 2008. A successful community-level strategy for conservation prioritization. *Journal of Applied Ecology*, 45: 1436-1445.

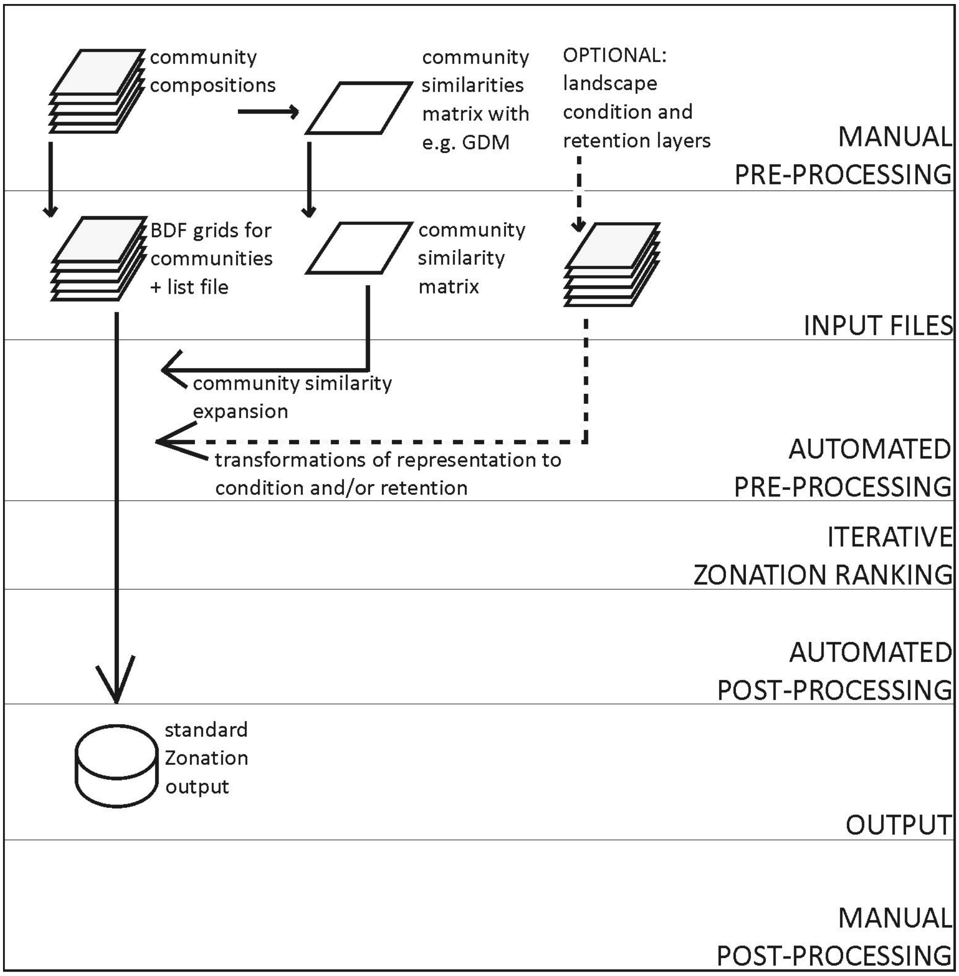
An example of implementation in Zonation in

Leathwick, J.R., A. Moilanen, S. Ferrier and K. Julian. 2010. Complementarity-based conservation prioritization using a community classification, and its application to riverine ecosystems. *Biological Conservation*, 143, 984-991.

Please note that Leathwick et al. utilize Zonation v.2.0. for which the data needs to be pre-processed before running the analysis. Zonation v.3.0. does the pre-processing by itself.

The theory and algorithm behind the community level analysis is explained in section 2.8.

Process chart for the analysis


community compositions
community similarities matrix with e.g. GDM
OPTIONAL: landscape condition and retention layers
MANUAL PRE-PROCESSING
BDF grids for communities + list file
community similarity matrix
INPUT FILES
community similarity expansion
transformations of representation to condition and/or retention
AUTOMATED PRE-PROCESSING
ITERATIVE ZONATION RANKING
AUTOMATED POST-PROCESSING
standard Zonation output
OUTPUT
MANUAL POST-PROCESSING

A process chart of a community level analysis. Please note that only the compulsory and most commonly used optional analysis components are presented - you can combine different components according to your specific needs.

Pre-processing of inputs

A pre-processing step you often would take is fitting habitat suitability models to existing species data and creating spatial predictions of habitat suitability or occurrences of biodiversity features across the planning region.

For a community level analysis, you need a community similarity matrix assigning a similarity value to all possible pairs of community types. The matrix can be compiled, for example, utilizing a generalized dissimilarity model (GDM; see Ferrier 2002). The matrix can be compiled, for example, in R with the package 'gdm' http://www.biomaps.net.au/gdm/.

A meaningful method for assigning weights to community types is to weigh them by their species richness. Weights are assigned in the first column of the biodiversity feature list file.

If you include landscape condition and retention, please see section 5.3.5. for necessary pre-processing steps.

Input files

For a community level analysis, you need

- A set of community type grids (section 3.3.2.1.)
- Associated biodiversity feature list file (section 3.3.2.2.). Weight community types according to their species richness.
- A community similarity matrix (section 3.3.3.4.) describing the overlap in species composition between community types.
- A run settings file with appropriate settings (section 3.3.2.3.)

It is possible to include considerations of cost, landscape condition and retention or other analysis components, depending on the specific aims and considerations in your analysis.

Analysis stages and settings

To run a community level analysis, you need to adjust your run settings file to include

```
[Community analysis settings]
load similarity matrix = 1
community similarity matrix file =
                  my_community_similarity_matrix.txt (the name of
                  your similarity matrix file)
apply to representation = 1
```

Output and its interpretation

When examining the representation of community types, it is worth noting that the coverage for communities after similarity transformation may not be equivalent to the coverage of the original community types. There are several alternative ways to obtain representation curves for the original, unexpanded community types:

- Use solution loading (section 3.5.2.). Load the output rank file from your community level analysis and give the original community types as biodiversity features, but without the community similarity matrix.

- List your original community types twice into your biodiversity features list file, but assign them zero weights for the second ones. This way you will get the representation curves for them in the same run in which you use the similarity expansion. If you do this, it is convenient also to use a groups file (section 3.3.3.12.), where you assign the layers with zero weights into one output group (and the similarity-transformed layers in another). That way you get a summary of representation levels for the original data automatically as well.

Strengths and weaknesses and further considerations

Strengths: fewer features are needed for the community level analysis than when working with individual species.

In many countries, conservation legislation is based on habitat types, which can be seen as surrogates for community types. Note, though, that accounting for variation in richness and overlap in community composition greatly increases the accuracy of the analysis compared to setting priorities for habitat types only.

Consideration of landscape condition and retention is a natural component in this analysis.

Weakness: Consideration of connectivity in the community context is more abstract than in the individual species context.

Link to tutorial

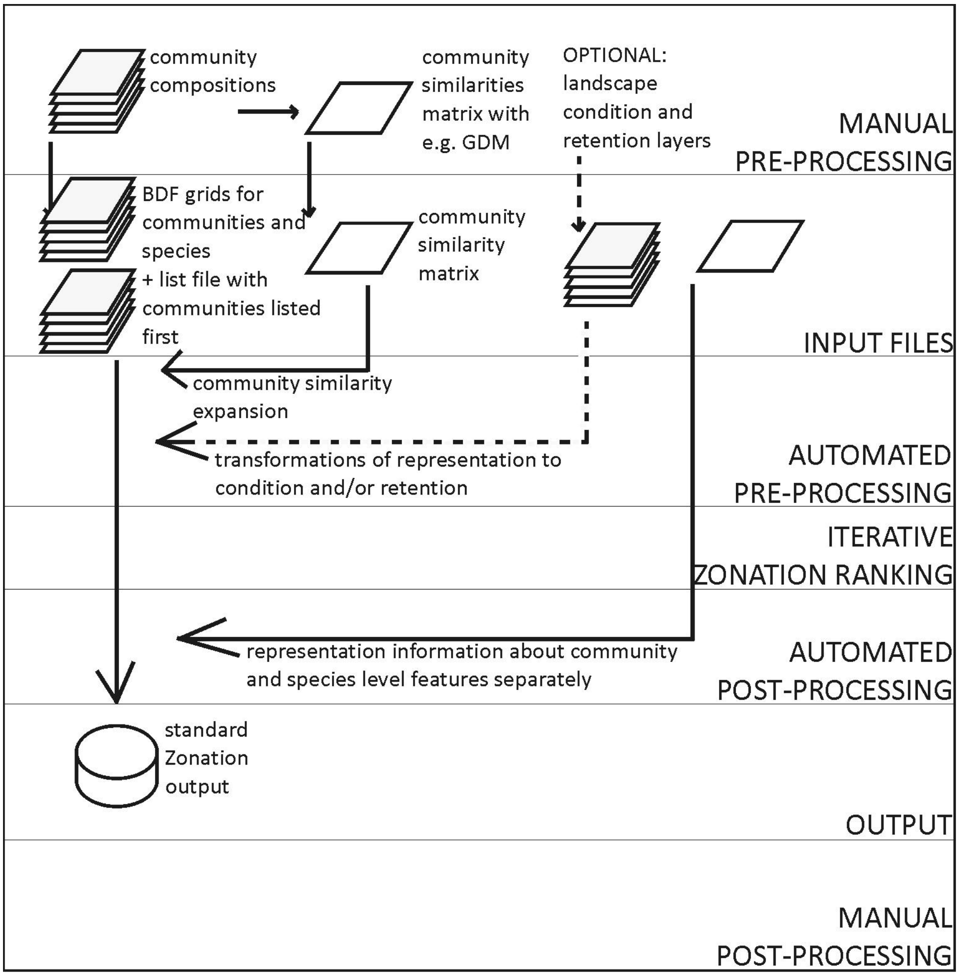
See Exercise 10 for an example of a community level analysis.

5.3.2 Combined community and species level analysis

Planning problem to be solved

Identifying conservation priorities based on a mixture of community types and occurrences of individual species. This approach may be useful when conservation decisions are mostly based on community types, but representation of some individual species needs to be considered on top of that.

Process chart of the analysis


community compositions
community similarities matrix with e.g. GDM
OPTIONAL: landscape condition and retention layers
MANUAL PRE-PROCESSING
BDF grids for communities and species + list file with communities listed first
community similarity matrix
INPUT FILES
community similarity expansion
transformations of representation to condition and/or retention
AUTOMATED PRE-PROCESSING
ITERATIVE ZONATION RANKING
representation information about community and species level features separately
AUTOMATED POST-PROCESSING
standard Zonation output
OUTPUT
MANUAL POST-PROCESSING

A process chart of a combined community and species level analysis. Please note that only the compulsory and most commonly used optional analysis components are presented - you can combine different components according to your specific needs.

Input data

To run a combined community and species level analysis, you need

- A set of community type grids (section 3.3.2.1.)
- A set of species occurrence grids (section 3.3.2.1.)
- A biodiversity feature list file (section 3.3.2.2.), in which you first list your community features and place individual species to the end of the list. This is necessary, as the community similarity matrix reads the features from the beginning of the file.
- A community similarity matrix (section 3.3.3.4.) describing the overlap in species composition between community types.
- A groups file (section 3.3.3.12.) to get representation curves separately for community and species level features. Assign your community level features to output group 1 and species level features to output group 2 by typing the respective values to column 1 of your groups file.
- A run settings file with appropriate settings (section 3.3.2.3.)

Depending on the specific aims and details of your analysis, you may want to include other input files, such as uncertainty layers, cost layer, landscape condition and retention layers etc.

Analysis stages and settings

Please refer to the previous section (5.3.1.) for necessary pre-processing and community level analysis settings.

To assign different types of biodiversity features to output groups, adjust your run settings file by including the following lines:

```
use groups = 1
groups file = my_groups_file.txt
```

under the `[Settings]` header (whereas the settings for a community level analysis would be under the `[Community level analysis]` header).

Link to tutorial

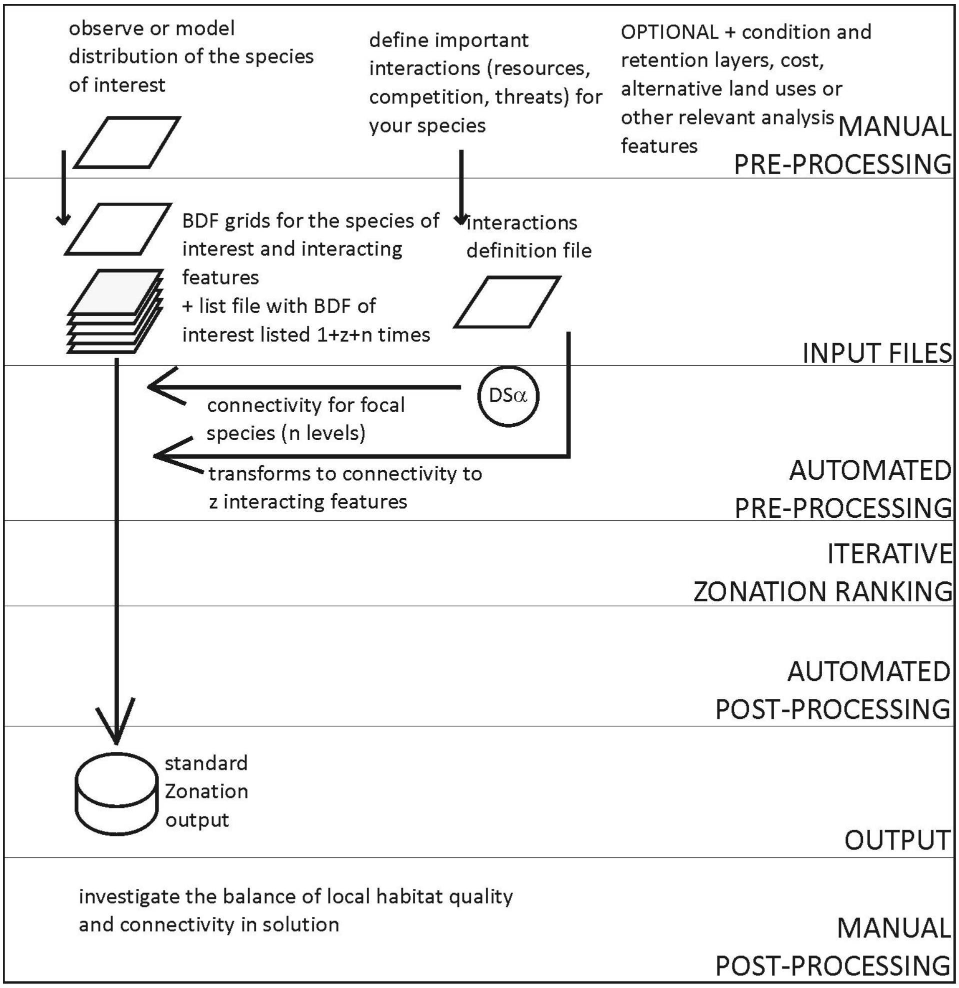
An example of combined community and species level analysis is provided in Exercise 10.

5.3.3 Single-species prioritization

Planning problem to be solved

Identifying conservation priorities for one (or a few) species with ecologically sound considerations of multiple aspects. Here, the aim is to optimize the protected area network to cover well-connected high quality patches for e.g. foraging and breeding, while avoiding harmful features such as human habitation or invasive species. Availability of resources and other vital species interactions can also be included. This requires ecological information about the species on top of occurrence data. Connectivity of multiple features and on multiple spatial scales can be accounted for.

Process chart for the analysis


observe or model distribution of the species of interest
define important interactions (resources, competition, threats) for your species
OPTIONAL + condition and retention layers, cost, alternative land uses or other relevant analysis features
MANUAL PRE-PROCESSING
BDF grids for the species of interest and interacting features
+ list file with BDF of interest listed 1+z+n times
interactions definition file
INPUT FILES
DSα
connectivity for focal species (n levels)
transforms to connectivity to z interacting features
AUTOMATED PRE-PROCESSING
ITERATIVE ZONATION RANKING
AUTOMATED POST-PROCESSING
standard Zonation output
OUTPUT
investigate the balance of local habitat quality and connectivity in solution
MANUAL POST-PROCESSING

A process chart of a single-species prioritization analysis. Please note that only the compulsory and most commonly used optional analysis components are presented - you can combine different components according to your specific needs.

Examples from literature

Rayfield, B., Moilanen, A. and Fortin M.-J. 2009. Incorporating consumer-resource spatial interactions in reserve design. *Ecological Modelling*, 220: 725-733.

Pre-processing of inputs

To be able to consider multiple ecological aspects of a species in conservation planning context, all the aspects need to be quantified and set into a spatial context. Considering connectivity within and between populations, two (or more) widths of dispersal kernels need to be defined for your species. Distributions of resources and other vital interactions, as well as threatening features, need to be transformed into grid rasters.

Input files

To account for connectivity at multiple levels in your prioritization, you need

- A biodiversity feature grid layer (section 3.3.2.1.) describing habitat quality (or probability of occurrence) for your species
- A biodiversity feature list file (section 3.3.2.2.). Here, the number of times you need to list your species layer is the number of connectivity considerations and interactions you include + 1. Assign different values of the dispersal α for each level of connectivity (see section 5.1.3. for details).
- A run settings file with appropriate settings (section 3.3.2.3.)

To account for connectivity to resources or other vital species interactions, or to favor sites that are not connected to threatening features such as human habitation or invasive species, you need

- Occurrence grid layers for the interacting features
- Add the abovementioned layers to your biodiversity feature list file. Please note that if you don't wish to account for the resource as such in the prioritization, you should give a zero weight for the resource in your list file. This way, the resource occurrence will only be used to estimate the conservation value of any cell in terms of how well it is connected to cells with resources.
- An ecological interactions definition file (section 3.3.3.8.). Define resources as positive interactions (type 1) and threatening features as negative interactions (type 2).

It is also possible to include considerations of cost, landscape condition and retention, or other components. Please refer to the appropriate sections for details on these.

Analysis stages and settings

To include ecological interactions, you need to adjust your run settings file to include the following lines

```
use interactions = 1

interaction file = my_interactions.txt
```

To indicate that distribution smoothing is applied in your analysis, you need to adjust your batch file:

- Set the third last parameter of your command line call to 1 to indicate that distribution smoothing will be done.

- Give a factor for multiplying the **species-specific α -values** as the second last parameter in your call. α -values (species-specific widths of kernel) are in the second column of your biodiversity feature list file. The factor is useful if you are interested in running multiple solutions with, e.g. assuming several levels of dispersal capabilities, because the factor allows you to multiply all dispersal capabilities simultaneously. Thus you do not need to change the parameters manually in the species list file after each run. If you do not wish to multiply the α -values, set this factor to 1.

Output and its interpretation

To assess the conservation solution, investigate the balance between habitat quality and connectivity (see Rayfield et al. 2009). The point is that you would ideally wish to lose only little local quality for the sake of improved connectivity. You can explore the balance of features and connectivity by examining the respective performance curves for the features and their connectivity transformed versions.

Strengths and weaknesses and further considerations

Strength: By including multiple ecologically realistic spatial considerations, it is possible to (partially) account for the processes that support persistence of the species in the long term.

Weakness: Systematic conservation planning is, in general, a multi-species enterprise. Placing much emphasis on a single species is questionable in that sense.

It is by all means possible to include basic occurrence data for a large set of species and, on top of that, more detailed considerations for a few key species.

Link to tutorial

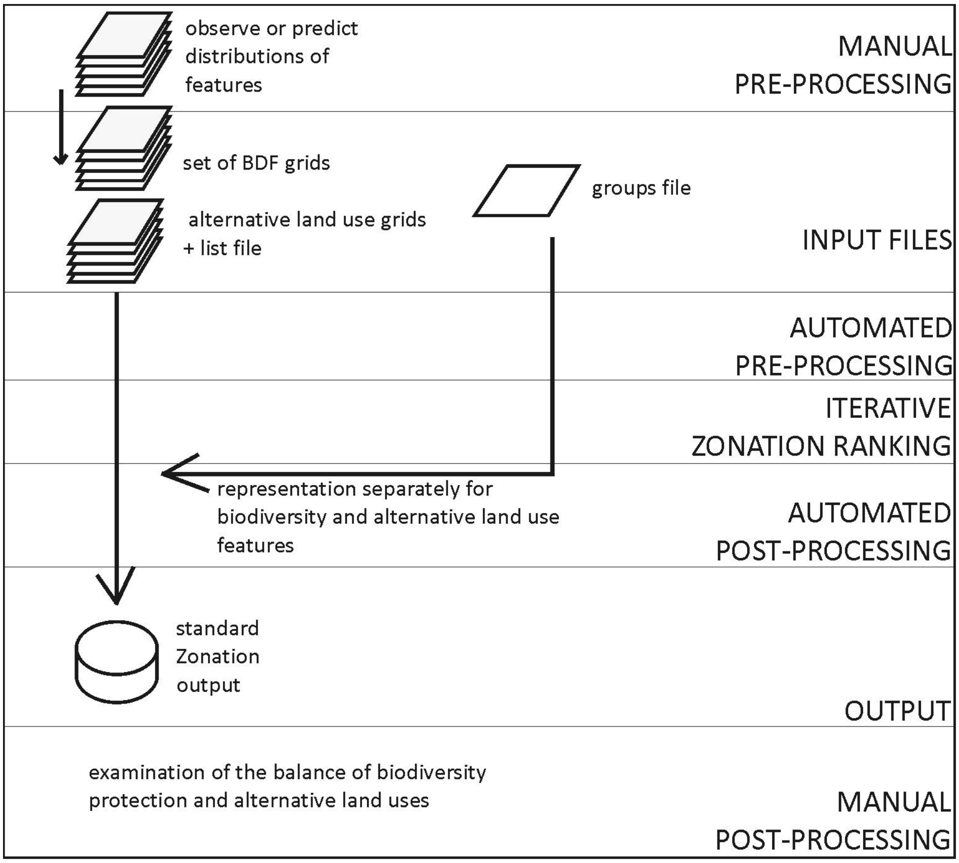
We do not provide a tutorial example for this analysis. Please refer to Rayfield et al. (2009) for a case example.

5.3.4 Balancing alternative land uses considering multiple costs

Planning problem to be solved

Finding a solution where land use is balanced between biodiversity conservation and other purposes. Here, the aim is to separate conservation priorities from competing land uses. Areas with ongoing or intended use for other purposes are given negative weights in feature weighting. Consequently, those areas will be removed fast in the ranking process. The analysis can help to identify multiple land use priorities and alleviate conflicts of interests.

Process chart for the analysis


observe or predict distributions of features
MANUAL PRE-PROCESSING
set of BDF grids
groups file
alternative land use grids + list file
INPUT FILES
AUTOMATED PRE-PROCESSING
ITERATIVE ZONATION RANKING
representation separately for biodiversity and alternative land use features
AUTOMATED POST-PROCESSING
standard Zonation output
OUTPUT
examination of the balance of biodiversity protection and alternative land uses
MANUAL POST-PROCESSING

A process chart of an analysis that balances multiple land uses. Please note that only the compulsory and most commonly used optional analysis components are presented - you can combine different components according to your specific needs.

Examples from literature

Moilanen, A., Anderson, B.J., Eigenbrod, F., Heinemeyer, A., Roy, D.B., Gillings, S., Armsworth, P.R., Gaston, K.J. and Thomas, C.D. 2011. Balancing alternative land uses in conservation prioritization. *Ecological Applications*, 21: 1419-1426.

Pre-processing of inputs

No pre-processing is needed for the alternative land uses analysis itself.

Other necessary pre-processing depends on the details of your analysis, for example, what kind of connectivity consideration you include.

Input files

To run an analysis for balancing alternative land uses, you need

- A set of biodiversity feature grids (section 3.3.2.1.)
- A set of alternative land uses grids (section 3.3.3.13.)
- A feature list file(section 3.3.2.2.). Please note that both biodiversity features and alternative land uses are listed in the same list file. The difference is that in the list file, alternative land use features are assigned negative weights.
- A groups file (section 3.3.3.12.) to get representation curves separately for the biodiversity features and for the negative features. Assign your biodiversity features to output group 1 and alternative land use features to output group 2 by typing the respective values to column 1 of your groups file.
- A run settings file with appropriate settings (section 3.3.2.3.)

Depending on the specific aims and details of your analysis, you may want to include other input files, such as uncertainty layers, cost layer, landscape condition and retention layers etc.

Analysis stages and settings

To run a simple balancing land uses analysis, you don't need any special settings other than the appropriate input files. See simple Zonation (section 5.1.1.) for the basic settings. The detailed combination of settings for your analysis depends on other features you want to include.

Adjust your run settings to include groups file

```
use groups = 1
groups file = my_groups_file.txt
```

Output and its interpretation

An optimal solution for balancing land uses would be one in which a small low-ranked fraction of the landscape includes most of the negative features, whereas a small top fraction would have high representation of the positive features (biodiversity). If such a solution can be obtained, then competing land uses and biodiversity conservation can be optimized simultaneously, without interest conflict.

Strengths and weaknesses and further considerations

Please note that the cost efficiency analysis and alternative land uses analysis are fundamentally different in the technical sense. In the cost efficiency analysis, biodiversity value of a cell is divided by cost. In alternative land uses analysis, the values of negative features are subtracted from the biodiversity value. Both of the approaches can be utilized in a single analysis, in which the overall structure becomes [(weighted aggregate of benefits) - (weighted aggregate of opportunity costs)] / (direct costs).

Link to tutorial

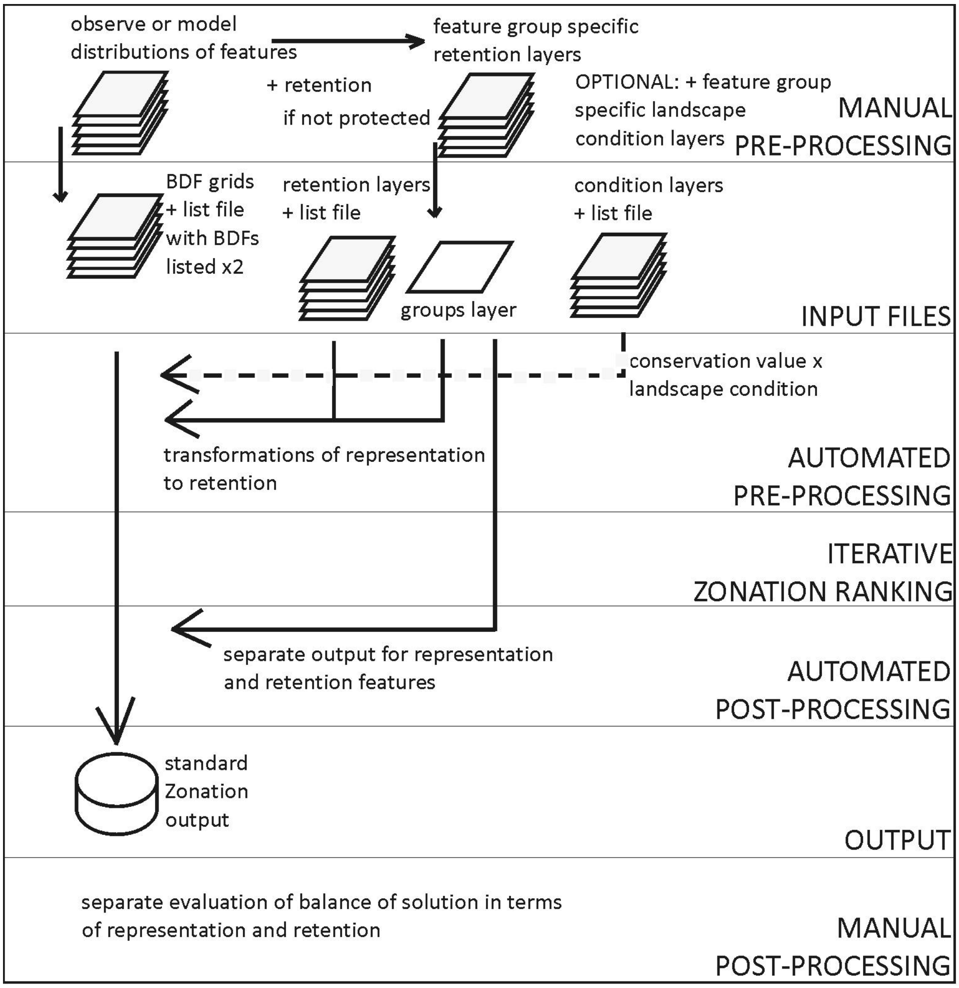
See exercise 2 for an example with one alternative land use layer.

5.3.5 Balancing representation and retention

Planning problem to be solved

Identifying conservation priorities when also the non-protected landscape can support some biodiversity value or when conservation value (population sizes, habitat condition) can even be increased with appropriate management. This analysis balances representation of features in the protected area network and their retention in across the full landscape. This approach also indirectly acknowledges that some species or communities are more in need of protection through protected area assignment than others. For example, species that are confined to old-growth forests may need protected areas more than generalist species of semi-open areas.

Process chart for the analysis


observe or model distributions of features
feature group specific retention layers
+ retention if not protected
OPTIONAL: + feature group specific landscape condition layers
MANUAL PRE-PROCESSING
BDF grids + list file with BDFs listed x2
retention layers + list file
condition layers + list file
groups layer
INPUT FILES
conservation value x landscape condition
transformations of representation to retention
AUTOMATED PRE-PROCESSING
ITERATIVE ZONATION RANKING
separate output for representation and retention features
AUTOMATED POST-PROCESSING
standard Zonation output
OUTPUT
separate evaluation of balance of solution in terms of representation and retention
MANUAL POST-PROCESSING

A process chart of an analysis that balances representation and retention. Please note that only the compulsory and most commonly used optional analysis components are presented - you can combine different components according to your specific needs.

Examples from literature

Moilanen, A., Leathwick, J.R. and Quinn, J.M. 2011. Spatial prioritization of conservation management. *Conservation Letters*, 4: 383-393.

Pre-processing of inputs

If you are running your analysis on community features, see section 3.3.3.4. about setting up a community similarity matrix.

For running an analysis with retention, you need to duplicate the list of features in your biodiversity features list file. The first set of layers are used to model conservation value with representation in reserves, the second set is used for modelling retention in the landscape without conservation action (or through a management intervention, using mode 2).

Several groupings of biodiversity features are necessary for including landscape (condition and) retention. First, it is advisable to divide the two copies of the biodiversity list into separate **output groups**. That way you can also compare the difference between priorities for representation and retention. Next, divide your biodiversity features into groups according to how past and projected changes in the landscape will affect **condition and retention** from their perspective. For example, for forest species the landscape condition would be decreased as a result from past clear-cutting of a patch, whereas for species that benefit from openings in the landscape a moderate clear-cut would have no effect on the condition. The same applies to retention. The groups do not need to be the same for condition and retention, although it would often make most sense. After grouping the features, you can set up a groups file (section 3.3.3.12.).

If you want to include condition: Create condition layers, one for each condition group. Assign each grid cell a value between 0 and 1 according to how past changes have affected the condition of that cell. Make a list of the layers into a condition layer list file. Please make sure that the rows in your condition layer list file match those in the groups file (see section 3.3.3.14.).

To create the retention layers, you need an estimate of how changes in the landscape will reduce (or improve) the conservation value of each grid cell for each retention group. Create retention layers and a list file (see section 3.3.3.15. for details). Please make sure that the rows in your retention layer list file match those in the groups file (see section 3.3.3.12.).Define retention mode (whether the quality conservation value is expected to increase or decrease for each retention group in the groups file as well.

Decide how you want to balance representation in selected reserves and retention in the landscape in your analysis. Crystallize the balance by assigning a value for the parameter β, which is a decimal multiplier for the retention layer weights for balancing between representation and retention.

Input files

For a balanced representation and retention analysis, you need

- A set of biodiversity feature grid layers (section 3.3.2.1.).
- A biodiversity feature list file (section 3.3.2.2.). List your biodiversity features twice: The first copy of layers models representation, the second copy will be modified to model loss if the cell is not selected.
- A community similarity matrix, if your features are communities (section 3.3.3.4.)
- A groups file assigning your features to retention groups and determining the retention mode (section 3.3.3.12.). It is also advisable to assign the layers used to model representation and the ones used to model retention to separate output groups. Insert suitable output group numbers to layers in the first column of this file.
- A set of retention layers, one for each retention group (section 3.3.3.15.)
- A retention layer list file (section 3.3.3.15.)
- A run settings file with appropriate settings (section 3.3.2.3.)

Depending on the specific aims and details of your analysis, you may want to include a set of condition layers, a cost layer, layers for alternative land uses etc.

Analysis stages and settings

To run an analysis to balance representation and retention, you need to adjust your run settings file to include

```
use groups = 1
groups file = my_groups_file.txt
use retention layer = 1
retention file = my_retention_file.txt
retention layers relative weight = value for β, a decimal multiplier for the
                                   retention layer weights for balancing between representation and
                                   retention
```

Output and its interpretation

When interpreting the output from retention analysis, it is important to note that Zonation tries to cover proportions of distributions, not absolute amounts. Retention layers are also treated as normalized distributions. A proportion of a small absolute amount (absolute fraction lost) may not have the same relevance as the same proportion from a large absolute loss from another feature.

Strengths and weaknesses and further considerations

Strength: An approach where landscape retention is considered alongside with representativeness of the protected area network is more realistic than considering the protected areas as the only fractions of landscape supporting biodiversity values. It is relevant to account for what would be lost in the absence of action.

Weakness: In the analysis, landscape dynamics and retention are dealt with in a relatively simplistic manner. Nevertheless, it is still a decent first approximation.

Link to tutorial

See exercise 11 for a tutorial example.

5.3.6 Habitat restoration and dynamic landscapes

Planning problem to be solved

Accounting for time aspect and habitat dynamics in conservation prioritization. The habitat restoration and dynamic landscapes analysis resembles the climate change setup (see section 5.3.7.) to an extent. The difference is that here distributions are expected to be dynamic, yet stationary, while in the climate context they are expected to be dynamic and non-stationary. Predictions of the distribution (habitat suitability) of species or other biodiversity features are produced for multiple time steps. The prevailing conditions at each time step are extracted from a scenario of habitat restoration actions at appropriate locations across the landscape.

The operational key to the analysis is inputting predicted habitat suitability for multiple species in a single analysis. Using core-area Zonation as cell removal rule will guide the solution to include suitable sites for all biodiversity features at every time step.

Examples from literature

Thomson, J.R., Moilanen, A., McNally, R., and Vesk, P. 2009. Where and when to revegetate: A quantitative method for scheduling landscape reconstruction. *Ecological Applications*, 19: 817-828.

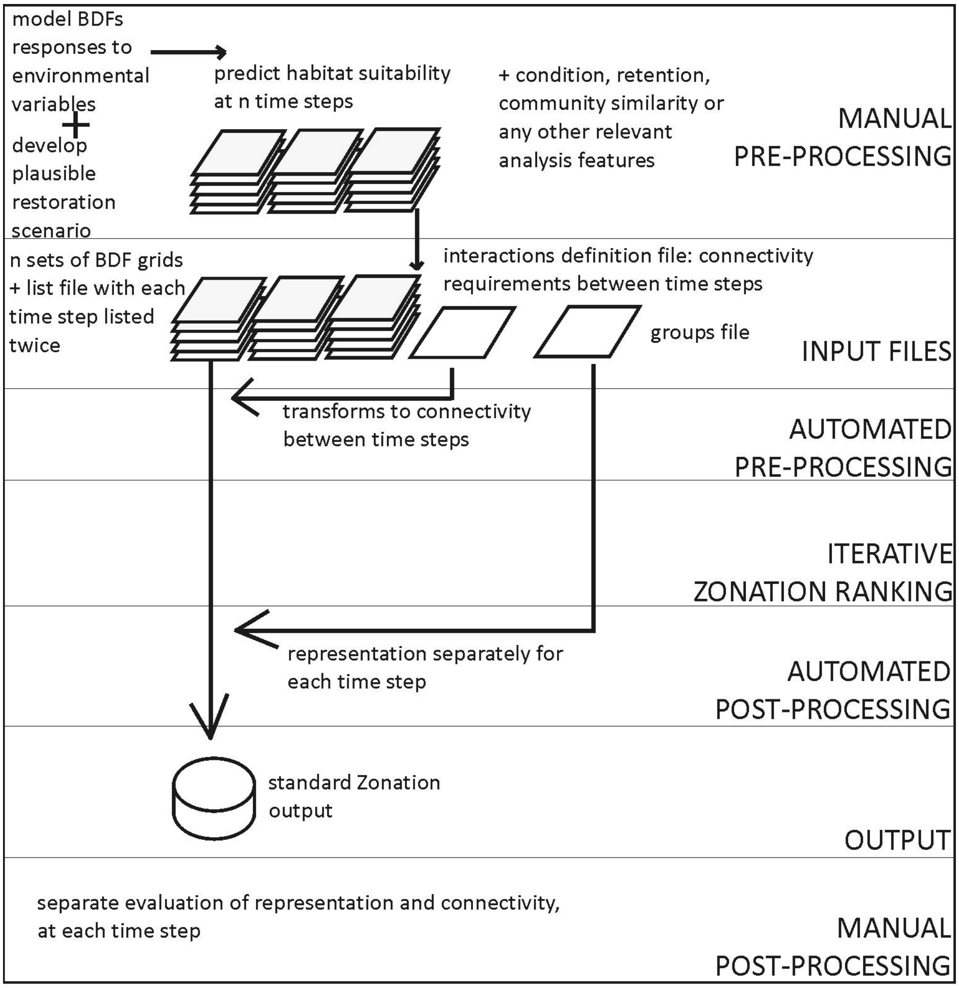
The theory and algorithm behind an analysis with dynamic landscapes is explained in section 2.11.

Pre-processing of inputs

A necessary pre-processing step in a dynamic landscape analysis is fitting habitat suitability models to existing species data and creating spatial predictions of habitat suitability or occurrences of biodiversity features across the planning region.

You also need to develop a scenario for habitat restoration or other landscape dynamics. In the scenario, understandably, restoration should be applied in locations where it is feasible, legal and sensible. If data about cost-efficiency of different actions is available, it can be utilized in the process. The scenario should be expressed as quantitative habitat properties for each grid cell at discrete time steps. The properties should be the most relevant ones explaining variation in species occurrences. Based on the scenario, use any distribution modeling approach to generate predictions of habitat suitability, usually expressed as probability of occurrence, for each of your species (or other biodiversity features) at each time step.

Process chart for the analysis


model BDFs responses to environmental variables
+
develop plausible restoration scenario
predict habitat suitability at n time steps
+ condition, retention, community similarity or any other relevant analysis features
MANUAL PRE-PROCESSING
n sets of BDF grids + list file with each time step listed twice
interactions definition file: connectivity requirements between time steps
groups file
INPUT FILES
transforms to connectivity between time steps
AUTOMATED PRE-PROCESSING
ITERATIVE ZONATION RANKING
representation separately for each time step
AUTOMATED POST-PROCESSING
standard Zonation output
OUTPUT
separate evaluation of representation and connectivity, at each time step
MANUAL POST-PROCESSING

A process chart of an analysis that considers landscape dynamics. Please note that only the compulsory and most commonly used optional analysis components are presented - you can combine different components according to your specific needs.

Input files

For a habitat restoration and dynamic landscapes analysis, you need

- Sets of biodiversity feature grids, one set for each time step (section 3.3.2.1.)
- A biodiversity feature list file (section 3.3.2.2.). Feature grids for all time steps are listed in the same file.
- A run settings file with appropriate settings (section 3.3.2.3.)
- A groups file (section 3.3.3.12.) to get representation curves separately for each

time step. Insert suitable values to column 1 of your groups file.

Optional files to be included in the analysis include

- A removal mask layer (section 3.3.3.9.) for masking in present good quality habitats that are not likely to be affected by habitat restoration. Using area inclusion mask is advisable in situations where good quality habitat is currently existing in the landscape.

Depending on the specific aims and details of your analysis, you may want to include other input files, such as uncertainty layers, cost layer, landscape condition and retention layers etc.

Analysis stages and settings

Use the predicted probabilities of occurrence for multiple time steps as a basis for conservation prioritization. Grid layers for all the time steps are listed in a single biodiversity feature list file. Adjust your run settings as running simple Zonation (section 5.1.1.) evaluating existing conservation areas, or as when expanding conservation areas (section 5.2.3.) when you are using a removal mask layer.

Adjust your run settings to include groups file

```
use groups = 1
groups file = my_groups_file.txt
```

Applying distribution smoothing, boundary quality penalty or other connectivity consideration in the analysis is advisable. Please refer to the corresponding sections for necessary input files and settings.

Output and its interpretation

The ranking in this analysis should be interpreted as a prioritization order that maximizes the availability of suitable habitat across all time steps **given that** the habitat restoration scenario created at stage 1. is successfully executed.

Strengths and weaknesses and further considerations

Weakness: As this analysis requires quantitative estimates of the habitat quality development, it is quite data intensive and complicated to be set up properly. The underlying scenario is susceptible to uncertainties at least two levels: the full implementation of the scenario complicated analysis may be uncertain, as are the real effects of action for habitat quality.

Connectivity between time steps could be set up in a multitude of ways. For example, the interlinkedness of the different time steps for one species can be emphasized by inputting the time steps as ecological interactions (see section 5.1.9.).

Link to tutorial

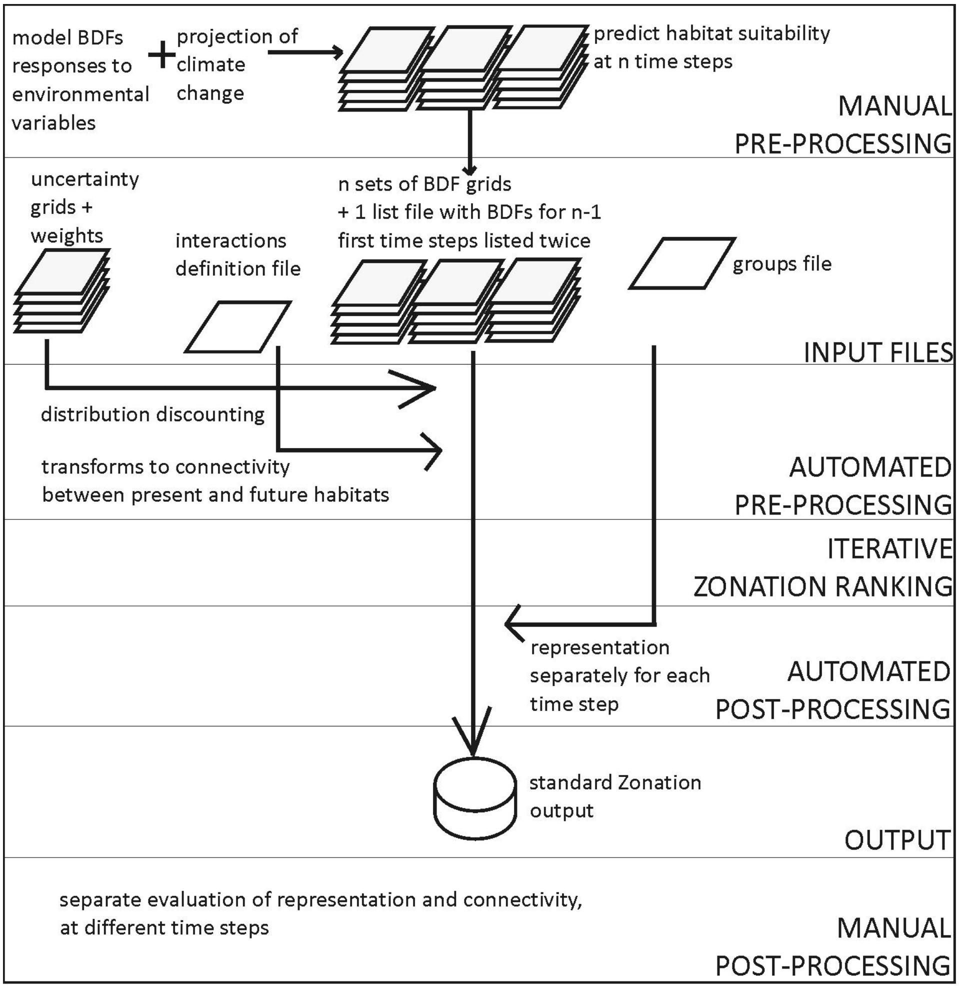
As the analysis is relatively simple to perform in Zonation, there is no tutorial for it. Please refer to tutorial 1 for selecting conservation areas or tutorial 7 if you are using an area inclusion mask.

5.3.7 Setups for climate change

Planning problem to be solved

Identifying priorities for conservation to account for changing distributions due to climate change. Connectivity of suitable habitats at different times needs special attention in this case. As pronounced uncertainties are related to considerations of climate change, they need to be addressed as well.

Process chart for the analysis


model BDFs responses to environmental variables
+
projection of climate change
predict habitat suitability at n time steps
MANUAL PRE-PROCESSING
uncertainty grids + weights
interactions definition file
n sets of BDF grids + 1 list file with BDFs for n-1 first time steps listed twice
groups file
INPUT FILES
distribution discounting
transforms to connectivity between present and future habitats
AUTOMATED PRE-PROCESSING
ITERATIVE ZONATION RANKING
representation separately for each time step
AUTOMATED POST-PROCESSING
standard Zonation output
OUTPUT
separate evaluation of representation and connectivity, at different time steps
MANUAL POST-PROCESSING

A process chart of an analysis that considers habitat quality and connectivity between time steps in the course of climate change. Please note that only the compulsory and most commonly used optional analysis components are presented - you can combine different components according to your specific needs.

Examples from literature

Carroll, C., Moilanen, A., and Dunk, J. 2010. Designing multi-species reserve networks for resilience to climate change: priority areas for spotted owl and localized endemics in the pacific North-West USA. *Global Change Biology*, 16: 891-904.

Pre-processing of inputs

Obtain a projection of climate development under a given emission scenario for one or more time steps. Please note that as there is large variation in climate projections depending on the climate model used and the underlying emission scenario, it is advisable to consider multiple models and scenarios. Create spatial predictions of future habitat suitability (or probability of occurrence) for your species or other biodiversity features. Again, as variation between different models is high, and model validation difficult or impossible, considering either a number of modeling strategies or using ensemble forecasting or similar strategy is advisable.

The connectivity between present and future suitable habitats are applied as ecological interactions (type 1; see section 3.3.3.8.). Define connectivity requirements for each species or other biodiversity feature are defined as widths of dispersal kernels (see section 5.1.3.)

To be able to account for uncertainty, you need to produce maps of uncertainty for each species or other feature. Here, you can utilize the standard deviation of your model predictions (see section 2.5.1.).

If you wish to consider landscape condition and retention in your analysis, please refer to section 5.3.5. for necessary pre-processing and other settings.

Input files

For an analysis to account for climate change, you need

- A set of biodiversity feature grid layers describing present distributions (section 3.3.2.1.)
- A set of biodiversity feature grid layers describing future distributions (section 3.3.2.1.)
- A biodiversity feature list file (section 3.3.2.2.). List both present and future layers in a single list file. List present biodiversity features twice: one to account for present distributions and another to be transformed to connectivity between present and future distributions.
- An ecological interactions definitions file (section 3.3.3.8.); note that here the interaction is not actually between species, but between present and future distributions of a single species)
- A set of uncertainty map layers (section 3.3.3.7.)
- An uncertainty weights file (section 3.3.3.7.)

- A groups file (section 3.3.3.12.) to get representation curves separately for each time step. Insert suitable values to column 1 of your groups file.
- A run settings file with appropriate settings (section 3.3.2.3.)

Optional input files include a removal mask layer (section 3.3.3.9.) to force existing protected areas in the top fraction.

Depending on the specific aims and details of your analysis, you may want to include considerations of cost, landscape condition and retention etc.

Analysis stages and settings

To identify conservation priorities that are likely to have high conservation value both in the present and the future, you need to adjust your run settings as follows:

Include interactions into your analysis (section section 5.2.9.)

```
use interactions = 1
interaction file = my_interactions_definitions_file.txt
```

Include groups file

```
use groups = 1
groups file = my_groups_file.txt
```

Include distribution discounting into your analysis (section 5.2.8.)

```
[Info-gap settings]
use info-gap weights = 1
Info-gap weights file = my_IG_weights_file.txt
```

Determine the value to the uncertainty parameter α in your program call (the fourth last parameter) in the command line. α -value determines the horizon of uncertainty in the data and is usually unknown. Thus you need to test generate solutions with several α -values to determine how the spatial pattern behaves with increasing uncertainty.

If you want to include current reserves (expanding conservation areas), landscape condition and retention, cost or other components, please refer to the corresponding sections for setting details.

Output and its interpretation

As the network should maintain both present and future biodiversity, examination of the output to evaluate representation should be done separately for present and future occurrences of species or other biodiversity features.

Strengths and weaknesses and further considerations

Strength: Allows accounting for the ongoing change in climate, which is bound to affect species distributions.

Weakness: Pronounced uncertainty prevents the solutions from being very solid. It is advisable to consider multiple scenarios of climate development and responses. As validation of distribution models for the future is difficult or impossible, a strategy to reduce sensitivity of the solution to uncertainty is highly recommended.

Link to tutorial

For this analysis, we do not provide a tutorial example. Please refer to Carroll et al. 2010 for a case example.

5.3.8 Administrative units analysis

Planning problem to be solved

Setting conservation priorities over multiple administrative regions. While conservation decisions are taken at regional level, the population dynamics of species extend throughout their ranges. This analysis allows for balancing global and local conservation needs.

There are two variants of the analysis. They emphasize global and local considerations in different ways.

Mode 1 allows regional priorities (weights) for features to be given. While weights vary, representation is taken as global. Effectively, this analysis is as the basic Zonation analysis but with spatially variable weights.

Mode 2 is fundamentally different in that it requires representation of all features in each area. Effectively, there are different weights globally and locally, and representation is tracked separately both globally and in each subregion. Loss of conservation value is computed from 1+N components, where 1 is for global analysis and N is for N subregions.

Examples from literature

Moilanen, A. and A. Arponen. 2011. Administrative regions in conservation: Balancing local priorities with regional to global preferences in spatial planning. *Biological Conservation*, 144: 1719-1725.

Pre-processing of inputs

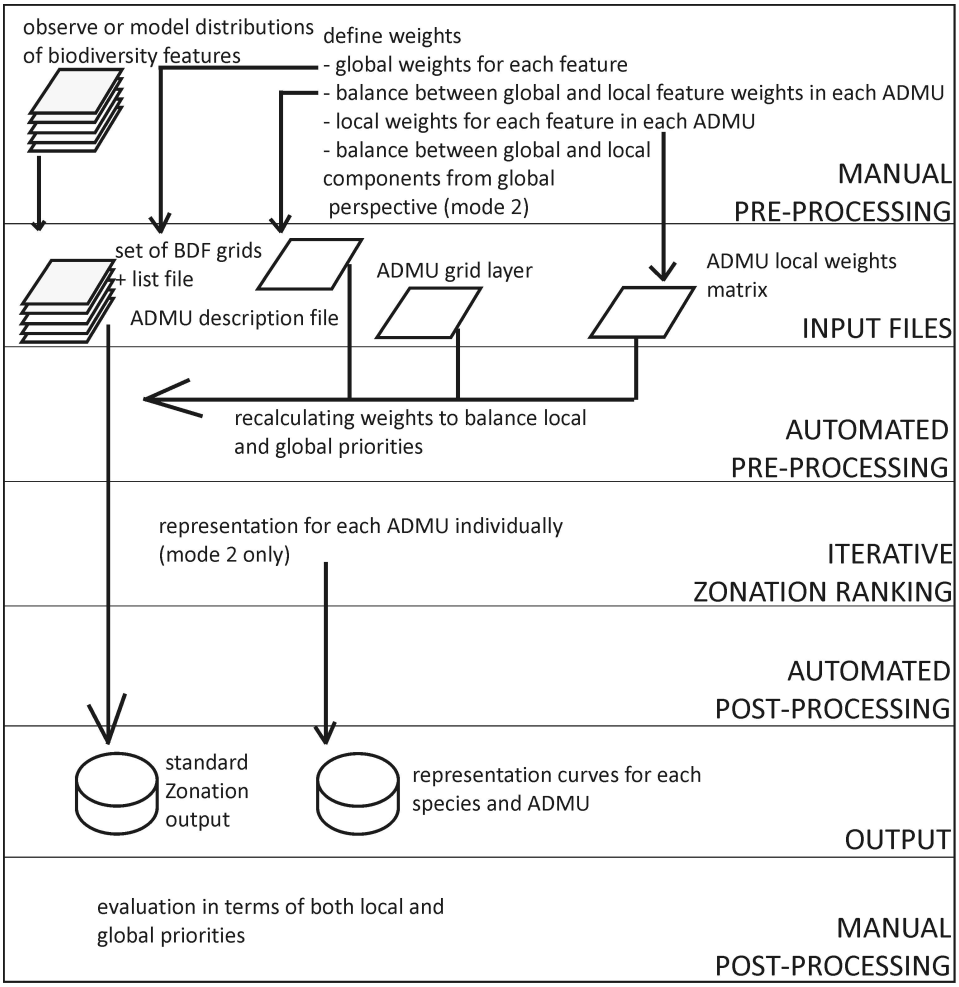
There are several weighting schemes that take place in the administrative units analysis. You need to assign

(i) Global feature weights for your species or other biodiversity features. Global feature weights go in to the weights column in the biodiversity feature list file.

(ii) Region-specific weights and parameters. Regional weights (G_A) determine the region's weight with respect to the other regions. The balancing parameter $ß_A$ describes how local and global feature weights are balanced in each administrative region. These are read in from the second and third column of your administrative units descriptions file.

(iii) Region-specific local weights for each of your species or other biodiversity features. These weights reflect purely local conservation priorities. These weights are compiled into a matrix where biodiversity features are in rows and administrative regions in columns. The order of the biodiversity features has to be the same as in the biodiversity feature list file.

Process chart for the analysis


observe or model distributions of biodiversity features
define weights
- global weights for each feature
- balance between global and local feature weights in each ADMU
- local weights for each feature in each ADMU
- balance between global and local components from global perspective (mode 2)
MANUAL PRE-PROCESSING
set of BDF grids + list file
ADMU description file
ADMU grid layer
ADMU local weights matrix
INPUT FILES
recalculating weights to balance local and global priorities
AUTOMATED PRE-PROCESSING
representation for each ADMU individually (mode 2 only)
ITERATIVE ZONATION RANKING
AUTOMATED POST-PROCESSING
standard Zonation output
representation curves for each species and ADMU
OUTPUT
evaluation in terms of both local and global priorities
MANUAL POST-PROCESSING

A process chart of an analysis that accounts for different priorities across several administrative regions. Please note that only the compulsory and most commonly used optional analysis components are presented - you can combine different components according to your specific needs.

Input files

To run an analysis over multiple administrative regions, you need

- A set of biodiversity feature grids (section 3.3.2.1.).
- A feature list file (section 3.3.2.2.). Assign the global weights for the biodiversity features in the first column of this file.
- An administrative unit map grid (section 3.3.3.16.)

- An administrative unit description file (section 3.3.3.16.). The region-specific weight should be in the second column of this file and the balancing parameter between local and global feature weights in the third column.

- A features by administrative units weights matrix (section 3.3.3.16.) defining local weights for biodiversity features in each administrative region.

- A run settings file with appropriate settings (section 3.3.2.3.)

Depending on the specific aims and details of your analysis, you may want to include other input files, such as uncertainty layers, cost layer, landscape condition and retention layers etc.

Analysis stages and settings

For an administrative units analysis, add the following lines to your run settings file (using your own file names of course):

```
[Administrative units]
use ADMUs = 1
ADMU descriptions file = my_ADMU_descriptions.txt
ADMU layer file = ADMUs_distribution_map.asc
ADMU weight matrix = ADMU_weights_matrix.txt
```

`calculate local weights from condition = 1` Presently inoperational dummy parameter.

`ADMU mode = 1` OR `2`, depending on whether you want to enforce local representation for all features (mode 2) or not (mode 1; see section 2.12.)

`Mode 2 global weight = 0.5` A parameter specifying the balancing of global and local conservation value when using ADMU mode = 2. This parameter is bounded between zero (only local considerations) and one (only global considerations influence value).

Output and its interpretation

All basic Zonation output, the priority rank map and the performance curves are produced. Additionally, there is a file that describes the occurrence levels of features in each administrative unit during landscape ranking (see section 3.4.2.).

Strengths and weaknesses and further considerations

Strength: This analysis can help especially in regional conservation priority setting to account for the global conservation status of species or other biodiversity features. Alternatively, it can be utilized to identify regional conservation priorities that contribute to the systematic conservation over a larger region.

Weakness: As multiple weighting schemes simultaneously take place on top of each other, the solution becomes sensitive to subjective weighting decisions. This can hardly be avoided and is not necessarily a problem, but is important to bear in mind while interpreting the output. Note that use of strong local priorities (mode 2) may result in loss of global efficiency because features may become protected also in regions where it is not (globally) efficient to do so. [Extra resources need to be expended where a globally common feature is locally rare.]

Link to tutorial

See tutorial exercise 9 for an example.

5.3.9 Setup combinations

There is a practically endless number of different analysis setups, and the setups described above only cover a small fraction of what is possible. We encourage creative use of Zonation analysis components. Bearing in mind a few considerations helps you develop successful and realistic analyses:

- Combine elements from different studies: species, communities, connectivity, costs, alternative land uses, condition, retention, etc. Choose the elements with respect to what kind of circumstances and limitations exist in your planning context.
- Develop complex analyses in stages. Start from simple. At each stage verify that the changes to the solution make sense following the addition of a new component to the analysis. This helps you trace back data problems or bugs in computations.
- When there are multiple plausible analysis setups, a selection frequency analysis can be done across the priority rankings to find out areas that (i) are always good, (ii) never are good, or (iii) are either good or not depending on settings. Areas of class (i) are most relevant for conservation.

Part VI

6 Tutorial & Examples

The purpose of this tutorial is to illustrate the use and function of different Zonation analyses. They also help you to familiarize yourself with the program and its settings. Later on, when working with your own data, you can use the example files to create your own input files.

The installation package includes the following example files:

For running basic Zonation
- species1.asc
- species2.asc
- species3.asc
- species4.asc
- species5.asc
- species6.asc
- species7.asc
- splist.spp
- splist_abf.spp
- splist_tbp.spp
- cost.asc
- set.dat
- set_cost.dat
- do_zig3.bat
- do_load.bat

For weighting species
- splist_w.spp
- splist_w_alu.spp
- do_w.bat

For using SSI species
- SSI_sp8.txt
- SSI_sp9.txt
- SSI_list.txt
- set_ssi.dat
- do_ssi.bat

For including distribution smoothing
- do_ds.bat

For including BQP
- BQPcurves.txt
- splist_bqp.spp
- set_bqp.dat
- do_bqp.bat

For including BLP
- set_blp.dat
- do_blp.bat

For including uncertainty analysis
- sp1_UC.asc
- sp2_UC.asc
- sp3_UC.asc
- sp4_UC.asc
- sp5_UC.asc
- sp6_UC.asc
- sp7_UC.asc
- UCweights.spp
- set_uc.dat
- do_uc.bat

For running the replacement cost analysis
- mask_rs.asc
- mask_towns_z3.asc
- set_maski.dat
- set_maske_z3.dat
- do_rs.bat
- do_towns.bat

For including directed connectivity (NQP)
- plu.asc
- rivers.txt
- NQPcurves.txt
- splist_nqp.spp
- set_nqp.dat
- do_nqp.bat

For community level analyses
- community1.asc
- community2.asc
- community3.asc
- community_similarity.txt
- groups_community_sp.txt
- bdlist_community.spp
- sp_community_list.spp
- set_community_similarity.dat
- set_community_sp.dat
- do_comm_nosimilarity.bat
- do_comm_similarity.bat
- do_comm_sp.bat

For condition and retention analyses
- condition1.asc
- condition2.asc
- cond_list.txt
- groups_condition.txt
- set_condition.dat
- do_condition.bat
- bdlist_retention.spp
- retention1.asc
- ret_list.txt

groups_retention_m1.txt
set_retention_m1.dat
do_retention_m1.bat

For considering administrative units
splist_ADMU.spp
ADMU_map.asc
ADMU_descriptions.txt
ADMU_weights.txt
set_ADMU.dat
do_ADMU.bat

Other files needed for tutorial and exercises
set_costds.dat
set_costbqp.dat
load_costds.bat
load_costbqp.bat
do_cost_ds.bat

As a final word, this tutorial does not include all variants of everything. When working with your own data, remember to carefully consider the cell removal rule you use. Analysis options that you might well wish to modify include edge removal, the warp factor, and adding of fake edge points.

6.1 Exercise 1

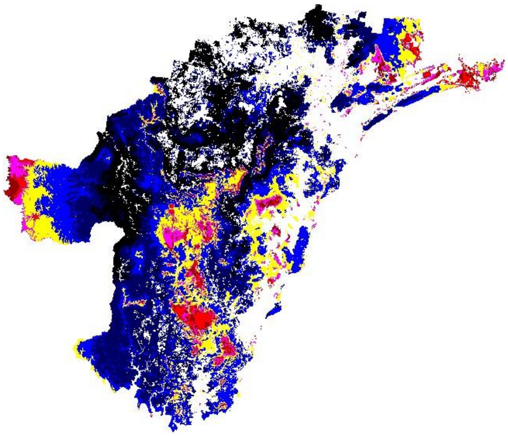
Getting started with the basic Zonation

Before starting see section 3.2. for how to operate the program either from the command prompt or the Windows GUI.

We start with a simple exercise by conducting the basic Zonation analysis. Let us think that there is a area in a remote country somewhere which is the home of seven rare species. We have been given a task to create a proposal for conservation network that will help to protect them. However, due to cost restraints the proposed conservation areas cannot be larger than 15% of the landscape. We decide to use the Zonation program to identify areas that have high priority for conservation. We also decide to use core-area Zonation as our planning method variant, because it best corresponds to our planning objectives (see section 2.3).

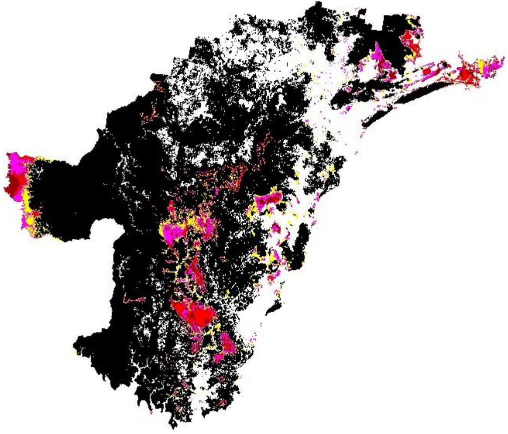
In your first species list file (**splist.spp**) you have a list of seven species distribution maps and species-specific parameters in front of the map names. Here all species are given an equal weight, but they have different dispersal α –values, as the use of surrounding landscape (e.g. home ranges) differs between the species. Because no other features are used in this first exercise, the last three parameters are given a dummy value of 1. In your first run settings file (**set.dat**) you can find all necessary parameters to run the basic core-area Zonation. Here we use edge removal, but you can also try to run this analysis without this option to see what kinds of effects it has to the result. The warp factor (here 100) has been adjusted so that the analysis runs smoothly but with fair accuracy. You can also test different warp factors to see how much of a difference it makes to the outcome and the running time - in our experience the difference in running time is substantial, but effects on accuracy are from minute to non-existent. You can run the batch file (**do_zig3.bat**) from Windows by double clicking it. (Note that the batch file is a text file that can be edited using notepad.) If you are using the Zonation GUI, you should open the Zonation project (the .bat file), and queue it to run - please see the section about the Zonation GUI for instructions. If you are using the "classic Zonation" color scale, the outcome of this analysis should look like this:

Batch file: do_zig3.bat
Rank = 1.0



As you remember, the biologically most valuable areas are shown here as red. Since we have the restriction of 15% on the area that we can protect, we now have to find those areas that are the best 15% of the landscape. You can identify these areas by using the color scale settings. Enter a color tab at 0.85 and give a color to it. Now the top 15% are shown by your color of choice (see GUI instructions). Right-clicking the image lets you save it into a graphics file, and the output of the analysis has been saved as a GIS raster that can be loaded to any GIS program for further processing. Please see descriptions of the .rank.asc and .prop.asc files.

Batch file: do_zig3.bat
Top 15% areas:



Remaining 30%:	Top 15%:
Area = 27 582	Area = 16 543
BL/A = 0.587	BL/A = 0.977

You can also identify areas that include at least, say, 30% of the distributions of all species. This corresponds to a minimum 30% target-based proportional coverage solution, which can thus be visualized for any method of aggregating conservation value, ABF, CAZ, TBF or GBF. To do so, open the "proportion lost" .prop.asc file and use the color sliders to show the top areas that include at least 30% for each species. (Somewhat confusingly, this is the color slider point 100-30=70 for the .prop.asc file.) Note that the proportional loss file links to the minimum species curve in the plot view - the red curve in the top-left panel. Below the figure, area is the number of grid cells in the chose top fraction, and BL/A is the boundary length to area ratio of the top fraction.

The cell ranking is only one part of the relevant Zonation output. Another part is a set of curves describing the absolute performance levels of species. The figure below shows the minimum (red line) and average (blue line) performance across all species for our basic Zonation analysis above. With this data set, a fraction of species distributions is lost already when only a small fraction of the landscape has been removed. This reflects the fact that the species in the sample data are both relatively widespread and that some of them have mostly non-overlapping distributions. (This is different from, e.g., the analysis of British butterflies, where the distributions of species were narrow and nested, and a substantial fraction of species distributions could be covered with 10% of the landscape. See Moilanen et al. 2005.) The data for generating these curves, and the respective species-specific curves, is output into a **.curves.txt file** under the name you have specified for your output.

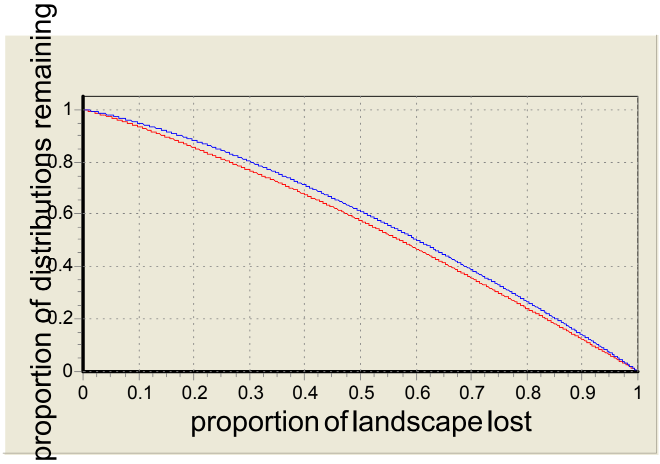


Figure showing how the average and minimum proportions of species distributions are declining as landscape is removed.

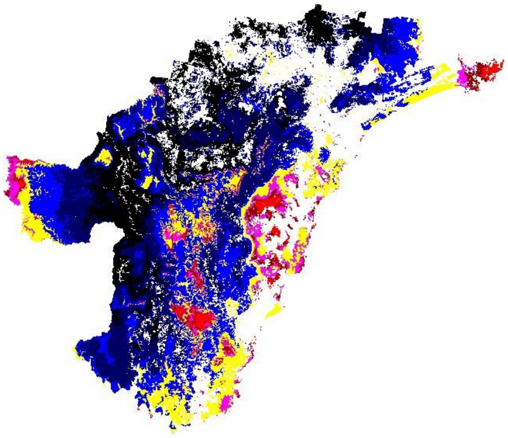
6.2 Exercise 2

Weighting of species or other biodiversity features

Now we have gotten started. However, two of our target species (species 2 and 3) are endemic and can not be found anywhere else in the world. Thus we want to enhance the protection of these species. This can be done by giving each of the species a weight of 2. The weights are determined in the species list file - thus we now use a different file called **splist_w.spp**, where the weights have been changed (naturally you could also use the same species list file as in Exercise 1 and just change the weights manually). Also the output files need to be renamed (**output_w.txt**) so that the program will not overwrite our earlier solution. The settings file used here is the same as in Exercise 1. Use the **do_w.bat** batch file to run the program or call the program yourself. Just remember to rename your output file so the program will not overwrite the solution from the previous exercise.

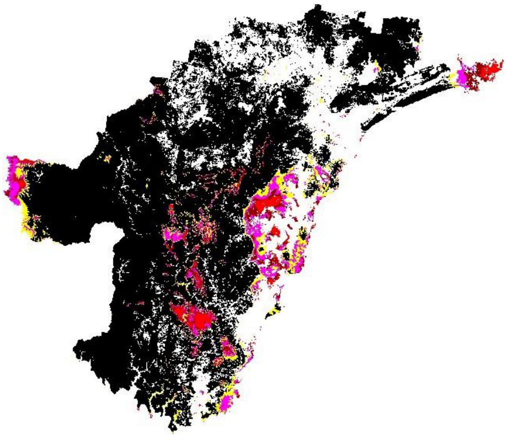
Batch file: do_w.bat

Rank = 1.0



Remaining 30%:
Area = 29 882
BL/A = 0.648

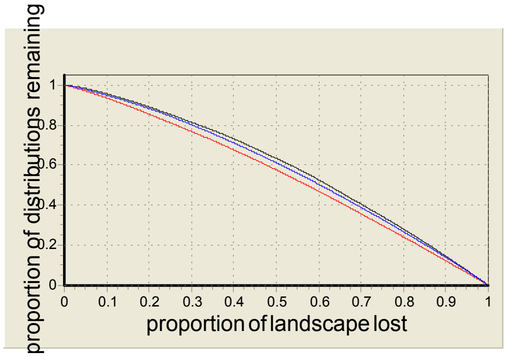
Rank = 0.15



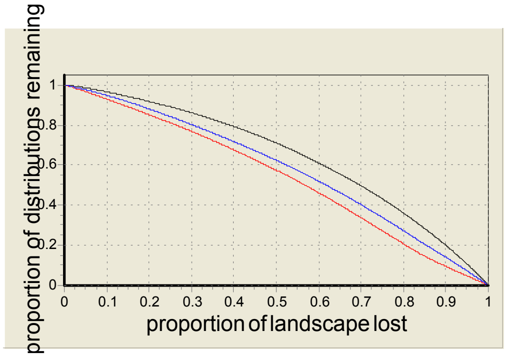
Top 15%:
Area = 16 543
BL/A = 0.932

As you see the weighting of species alters the spatial distribution of the highest value cells. Here more importance has been given to the areas where the two endemic species occur (e.g. west and south coast) and less to those areas that have a high representation for the other species (e.g. the peninsula on the east and the north-east region). A difference between the basic and weighted Zonation runs can be seen also in the species distributions curves in the Plot view -window.

Basic Zonation



Weighted Zonation



If you compare the two solutions, you can find that the relationship between landscape loss and **average** biodiversity protection for all species is quite similar in both cases (blue line). However, when species 2 and 3 are weighted, they retain a relatively higher proportion of their distributions through the cell removal process compared to the basic Zonation run where no weights are used. In turn the minimum proportion retained (species that has the lowest protection) is smaller when using species weighting. The two graphs above show the differences for species 2 (black line) between basic and weighted solution.

Solution comparison (see section 3.5.1.3) is a useful feature that can be used for visually comparing differences between two solutions. The figure below shows the overlap and differences in the best 15% of landscape between the basic and weighted Zonation runs. Overlapping regions are shown in yellow and areas that are in only one of the solutions are in green (dark green: no weights - light green: weighted).

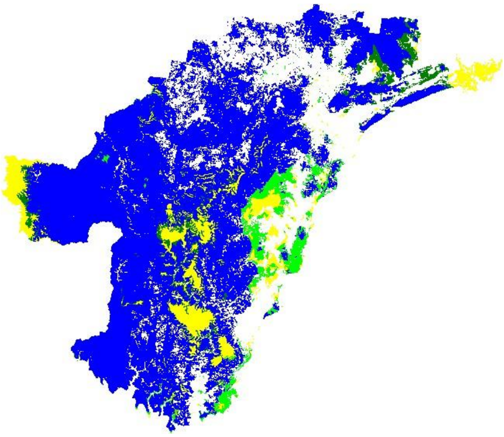


Figure showing differences in top 15% fraction between the non-weighted (dark green) and weighted (light green) solution.

Negative weights - accounting for alternative land uses or negative effects to biodiversity

In Zonation v.3.0, it is possible to include features that have a negative effect for conservation value. These can be, for example, areas that have high priority for urban construction, farmland, or other purposes. The analysis is implemented by treating the alternative land use layer as any biodiversity feature layer, but assigning it a negative weight in the biodiversity feature list file.

For the sake of example, let's imagine that the values in file **plu.asc** represent suitability of the grid cells for the expansion of urban development. The most suitable areas are likely to be zoned for urban construction in the near future, so we do not want to give them too much weight to alleviate the conflicting interests. We assign the layer a weight of -2.0 in a new species list file called **splist_w_alu.spp** which otherwise has similar settings to the list file we used in the previous example. Remember, again, to give a new name for your output files (**out_w_alu**) not to overwrite the outputs from previous runs. Run the analysis with the same settings file as previously, calling the program with the **do_w.bat** batch file.

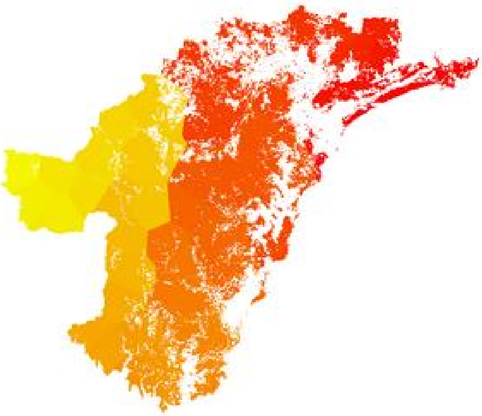
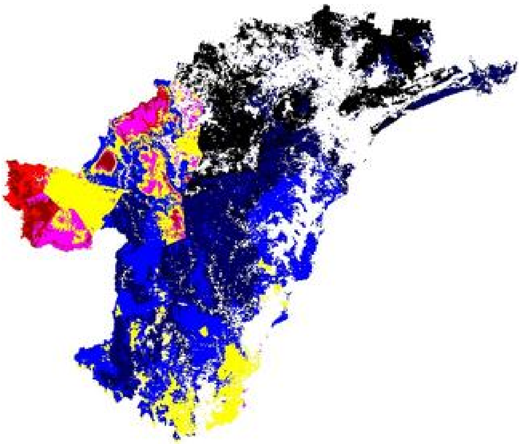


Figure showing suitability for expansion of urban development. Suitability increases with the intensity of red colour.

Batch file: do_w_alu.bat
Rank = 1.0

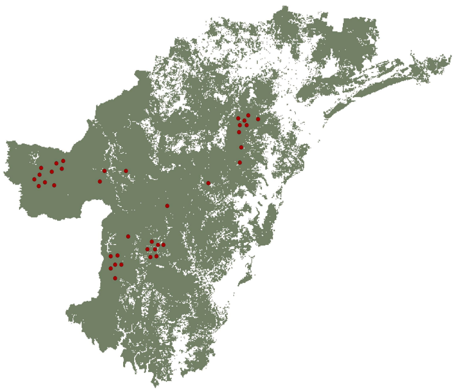


Compare the ranking map from this analysis to the one without the alternative land use layer. Sites that would have been high priorities for biodiversity were now removed from the solution early on, as they were assigned high values of a negative feature. Look at the representation curves in **.curves.txt** file to see how the negative weights affected species representation. When you look at the minimum representation curve, it is worthwhile to remember that it is the negative feature and not an actual biodiversity feature layer.

6.3 Exercise 3

Species of special interest

Let's assume that in addition to the seven species we work with so far, there is another two species which we should protect. However, these species are very rare and poorly known, and we only have a set of occurrence points indicating where the species have been observed. Unfortunately this data is not sufficient enough to allow us to model the species distribution to cover the entire study area. Yet, we want to include the species into our analyses and therefore enter them as species of special interest (SSI), simply using the point data we have. To do this we have created two text files, each of them listing the exact coordinates of the records for one species (**SSI_sp8.txt** and **SSI_sp9.txt**). We have also defined all the species-specific parameters in an additional SSI species list file (**SSI_list.txt**).



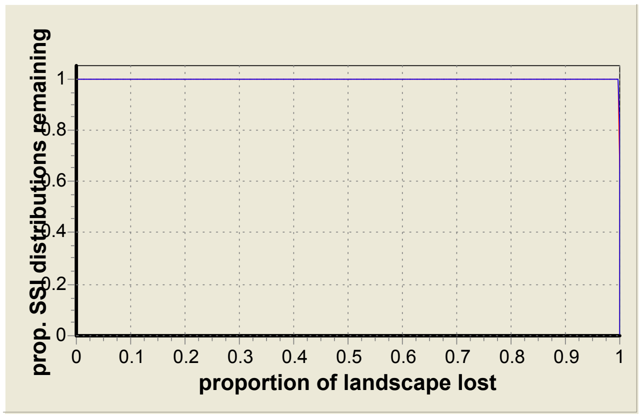
i) SSI species 1

ii) SSI species 2

Two maps showing the point occurrence data of our two SSI species, respectively.

The analysis is run as in Exercise 2, except that now we need to activate the SSI species option and give the name of our SSI species list file. You can either do this manually in windows interface or run the program from command prompt using the modified settings file (**set_SSI.dat**). You can call the program yourself or use the **do_ssi.bat** file to run the analysis.

As you can see, the inclusion of the two extra species brings hardly any changes to the results. This is because the points cover only a very small fraction of the study area, and it is therefore easy for the program to include them to the top fraction without altering the spatial distribution of high value cells in our solution. Also, because the area covered by the points is so small, the cells with SSI species receive very high values. Thus it follows that the full distribution (ie. all the points) of SSI species are practically always included to the top fraction. This is evident also from the SSI species graph in the species info window, from which we can see that the entire distributions for both SSI species are retained till the very end of the cell removal process.



Picture of SSI species curves in the Plots window.

Note that the same information that is displayed in the graph, is also outputted in .SSI_curves.txt -file, which the program produced during this analysis.

As the inclusion of the two SSI species contributes very little to the final result, from now on we keep them separate from the main analysis and continue with the original seven species.

6.4 Exercise 4

Adding aggregation into the analyses

Now we have identified sites that have a high occurrence of our target species (weighting the two endemic species). But the areas are quite fragmented, which is never a good quality in a reserve network. Thus we want to produce a more aggregated solution.

To do this, we try two different aggregation methods, **distribution smoothing** and the **boundary quality penalty**. Both methods favor the selection of contiguous cell groups, rather than selecting more fragmented sets of cells. This in turn offers advantages in terms of greater connectivity, and can also promote more practical and cost-effective management. Note that it is not recommended to use several aggregation methods simultaneously due to difficulties in interpreting the results.

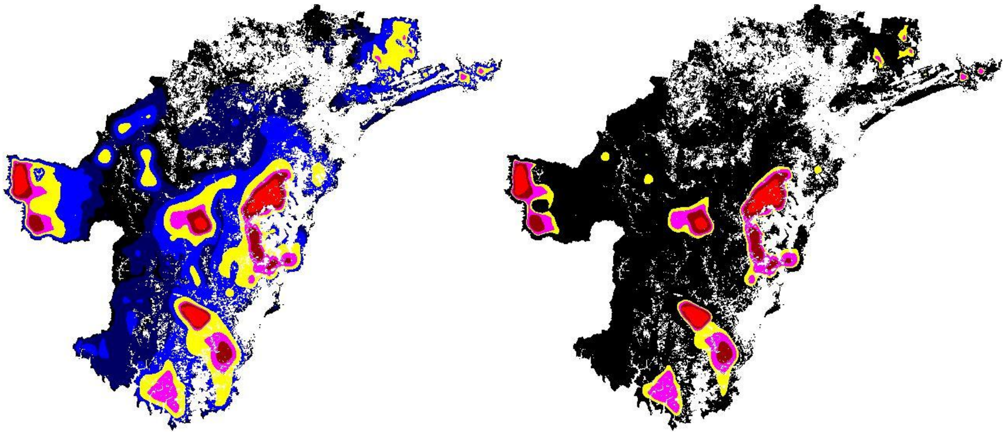
4a. Distribution smoothing

Using this method results more aggregated solutions based on the connectivity of sites. To add the distribution smoothing to the analysis we use the same input files as in Exercise 2, but we call the program with a new batch file, **do_ds.bat**. Note that the species-specific smoothing is defined in the species list file, where the width of the smoothing kernel for each species is given in the 2nd column, and in the call itself, where the third last parameter (value = 1) indicates that distribution smoothing will be done.

Batch file: do_ds.bat

Rank = 1.0

Rank = 0.15



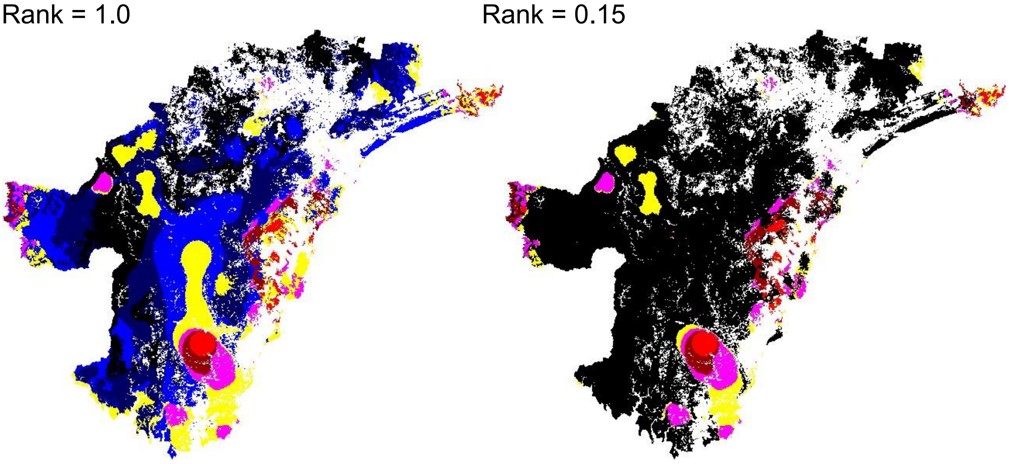
Remaining 30%:
Area = 27 493
BL/A = 0.205

Top 15%:
Area = 16 543
BL/A = 0.285

4b. Boundary quality penalty

The boundary quality penalty (BQP) calculates the most valuable sites based on both the value of the cell and the effects of habitat loss in the surrounding cells. The effects of fragmentation (loss of neighborhood cells) are species-specific and thus the BQP also takes in to account how different species are influenced by fragmentation and habitat loss. BQP analysis demands prolonged computation times compared to our previous method, because the loss of a cell now has an effect on occurrence levels in nearby cells, which has to be taken into account in calculations. To include the BQP to our analysis, we now use the settings file **set_bqp.dat**. Because BQP substantially increases computation time, the warp factor has been increased to 500 (in production runs we recommend warp factor of 100 or lower). We also need to use a new species list file, which gives the program all species-specific responses to fragmentation needed for running BQP. Thus use the **splist_bqp.spp** file as species list file. You can use the **do_bqp.bat** file to run the analysis or call the program yourself.

Batch file: do_bqp.bat

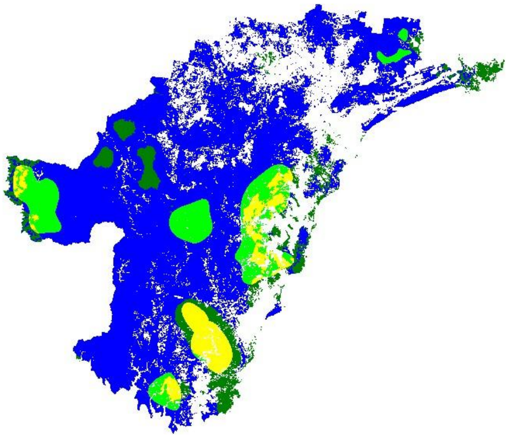

Rank = 1.0

Rank = 0.15

Remaining 30%:
Area = 30 875
BL/A = 0.563

Top 15%:
Area = 16 543
BL/A = 0.706

As you see, the two solution are fairly different. To get a better picture of the main differences use the solution comparison. You can either use the top 15% fraction of the two solutions for the comparison (see figure below), or you can check from the results how large fractions were needed to protect the 30% of all species distributions and use these fractions for the comparison.



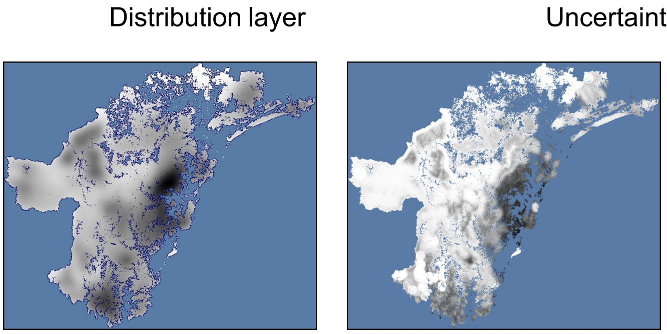
Comparison between top 15% areas selected by Zonation with distribution smoothing and Zonation using the BQP. Overlapping areas are shown in yellow, light green areas are present only in distribution smoothing solution and dark green areas only in BQP solution.

6.5 Exercise 5

Uncertainty in species information

Due to lack of data, uncertainties in species distributions are a common problem encountered in ecological studies. However, these uncertainties can be accounted for in reserve selection when using the Zonation program. Conceptually, the program uses uncertainty analysis to focus on sites where the prediction uncertainties are low compared to the predicted representation levels. Thus the program prioritizes sites that have both high abundance and low uncertainty.

To continue with our exercise, let us think that our species distribution data has been provided by statistical species distribution models. Models contain many uncertainties in their predictions and we want to take this into account when we are selecting the best sites to be protected. We have uncertainty layers for each of the species (**sp1_UC.asc**, **sp2_UC.asc** etc.), which show the spatial distribution of uncertainty in our data. The pair of figures below shows as an example the distribution and its error surface for one species.



Distribution layer

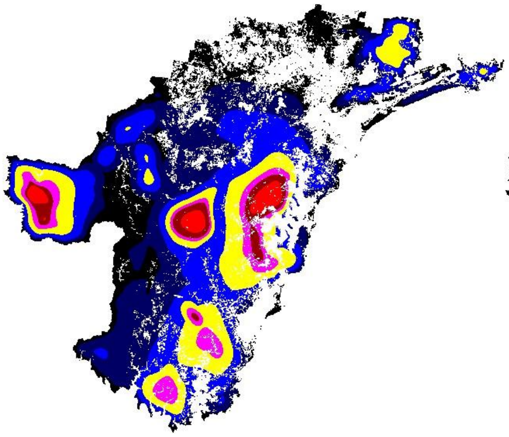
Uncertainty layer

Figure showing the distribution of species 2 and its corresponding uncertainty layer. White indicates high occupancy levels or high error, and black low occupancy or low error. You can display these maps by selecting the species from the tree in the Project view (for viewing the uncertainty layer remember to select the "Wmap" -option). Save by right-clicking on the image.

All uncertainty layers are listed in the **UCweights.spp** -file, where you can also find the species-specific uncertainty weights. In this exercise we want to stress the certainty of data equally for all species, thus we give them an identical weight of 1. We use the same species list file as in Exercise 4a, but the settings file needs to be changed to **set_uc.dat**. Again, use the **do_uc.bat** batch file to run the program or call the program yourself. In this exercise, where uncertainty analysis is done together with distribution smoothing, the program first calculates each cell a new value based on the uncertainties in the data and then uses these values for the aggregation part.

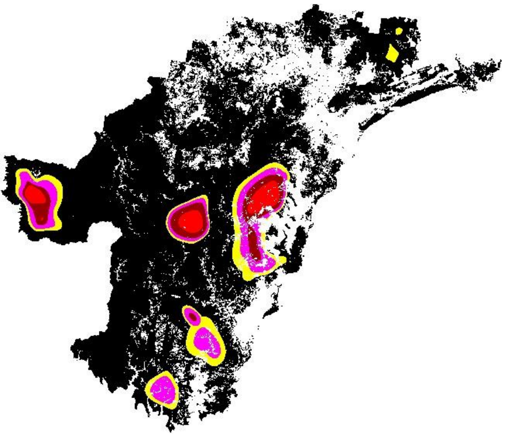
Batch file: do_uc.bat

Rank = 1.0



Remaining 30%:
Area = 28 193
BL/A = 0.194

Rank = 0.15

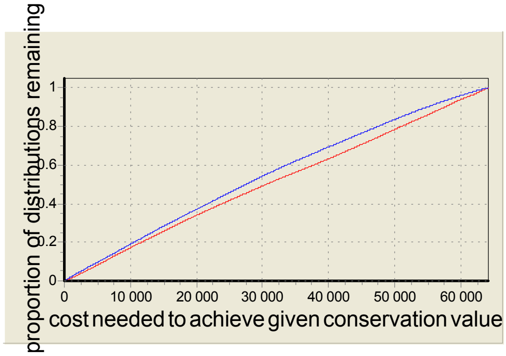


Top 15%:
Area = 16 543
BL/A = 0.273

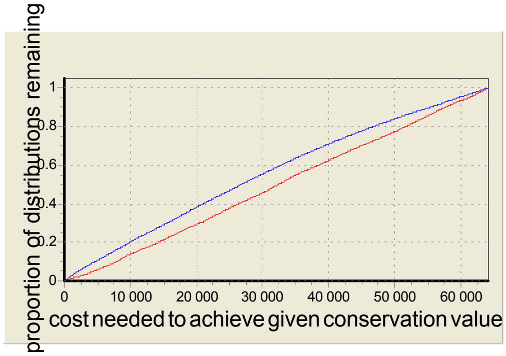
6.6 Exercise 6

What about the costs?

As in many cases in conservation biology, costs play a vital part in reserve planning. So far we have been looking for sites based solely on their biological value and ignoring the possible costs (except in terms of land area constrains). But, assume that in our study area land is most expensive at the south-west and cheapest at the north-east region. Now we need to determine how much the protection of our proposed areas would actually cost. To illustrate this we first load the two solutions calculated in Exercise 4 with a cost layer (**cost.asc**) and compare them. By loading a already produced solution the program can calculate the increase of cost as cells are removed in the exact same order as they were when producing the solutions in the Exercise 4. Use the two batch files, **load_costds.bat** and **load_costbqp.bat**, to see how large differences, if any, there is in the land costs between the two solutions.



Distribution smoothing

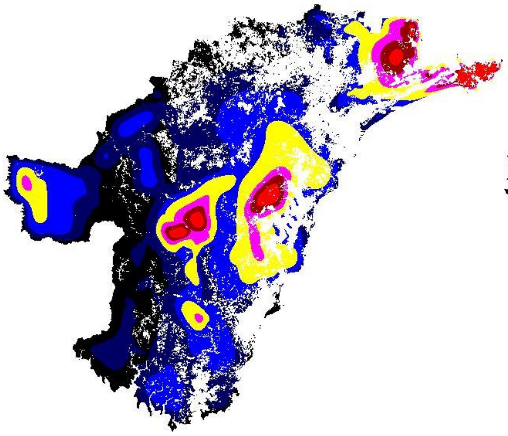


Boundary Quality Penalty

Another possibility would be to include the cost into the analyses during the cell removal process. This way the program would calculate a solution that has both high conservation value and low demands for resources. This is achieved by selecting cells that have a high conservation value/cost ratio. To do this we rerun the analysis from Exercise 5 using a cost layer (instead of loading it with the cost layer, as we did above). Again, we use distribution smoothing as our aggregation method, but you can choose other aggregation methods as well. The run settings for this exercise are defined in **set_costds.dat** -file. Use the **do_cost_ds.bat** -batch file to run the analysis. How do the most important areas change compared to the solution from Exercise 5? Compare also the species distribution curves of the two solutions, what changes do you see?

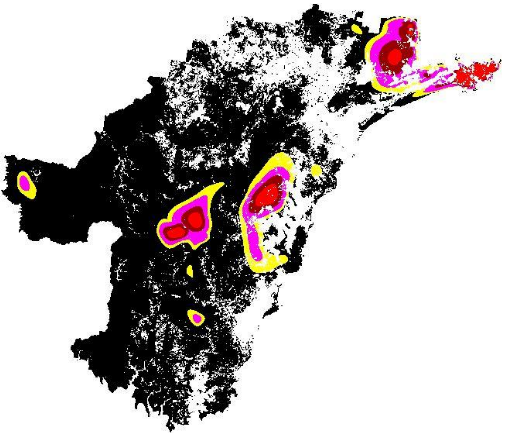
Batch file: do_cost_ds.bat

Rank = 1.0



Remaining 30%:
Area = 37 682
BL/A = 0.717

Rank = 0.15



Top 15%:
Area = 16 543
BL/A = 0.329

6.7 Exercise 7

What about the already existing reserves?

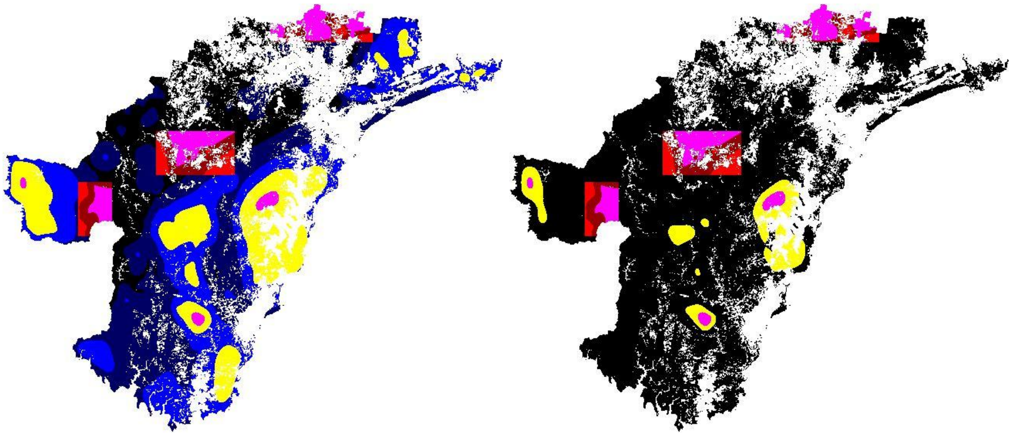
When it comes to reserve network planning, managers seldom have the chance to start from a fresh table. In many cases the target area already includes older reserves or areas that are ear-marked for other land uses such as agriculture, forestry or habitation etc. Thus, one often has to take into consideration areas that either can not be included or need to be included into the reserve network. Let us think that our example landscape already has a couple of reserves that we have to include to the final solution.

To do this we use the mask option which allows us to classify cells to different categories, which in turn define the cell removal order. In the tutorial package you can find a mask file which includes three reserves (**mask_rs.asc**). We run the analyses once more as in Exercise 5, but this time we include the mask (see **set_maski.dat**). Thus call the program with a new batch file **do_rs.bat**.

Batch file: do_rs.bat

Rank = 1.0

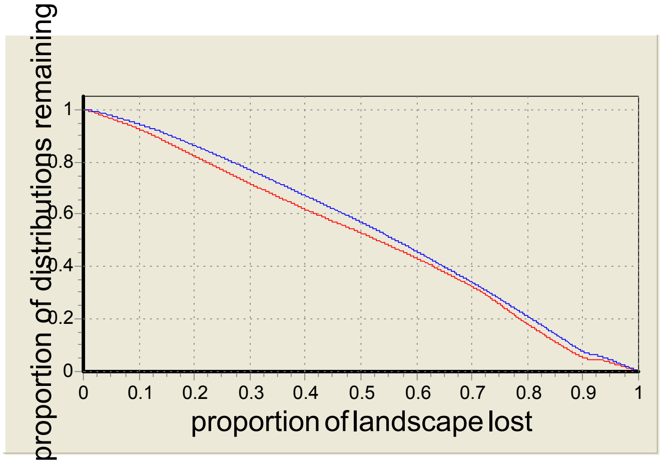
Rank = 0.15



Remaining 30%:
Area = 30 830
BL/A = 0.271

Top 15%:
Area = 16 543
BL/A = 0.369

Compare the solution to the results from Exercise 5. How has the representation of species altered now that the old reserves are included in the top 15% fraction? Note the differences at the end of the species distribution curves. As you see, the solution received with the old reserves is suboptimal, since the program is forced to include poor quality areas in the top fraction.

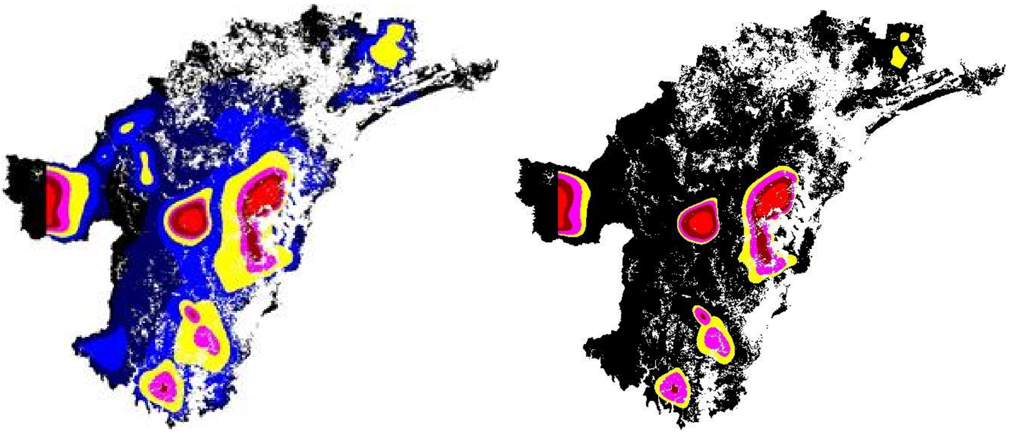


Average and minimum performances when old reserves are included to the top fraction.

In some cases the situation is the other way around, where areas can not be included to the reserve network for some reason. Thus these areas need to be first masked out from the landscape, before ranking the rest of the area. See how this kind of masking would change the results in our study area by running the analysis with the batch file **do_towns.bat**, where we have two imaginary regions marked as residential area (**mask_towns_z3.asc**). In this exercise we use the **set_maske_z3.dat** -settings file, all other input files are identical to the Exercise 5.

Batch file: do_towns.bat

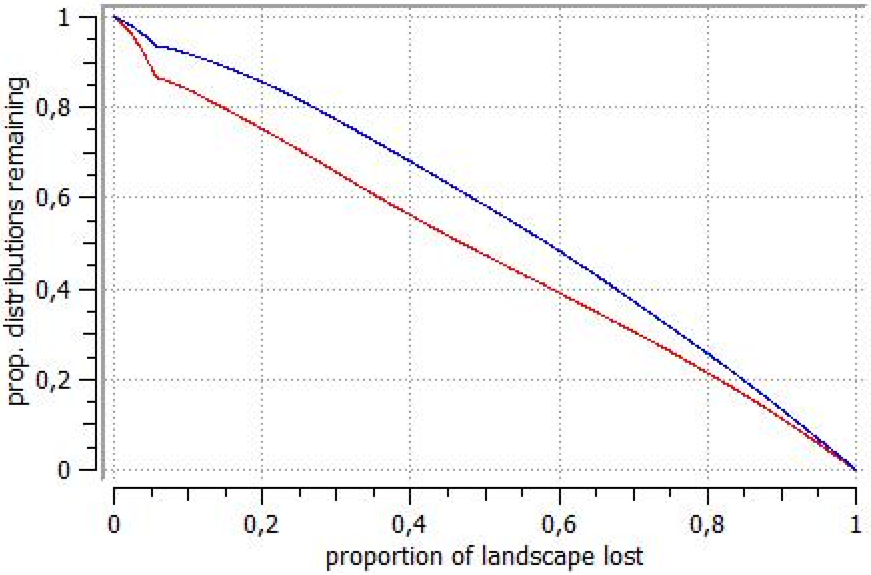
Rank = 1.0 Rank = 0.15



Remaining 30%:
Area = 32 360
BL/A = 0.251

Top 15%:
Area = 16 543
BL/A = 0.281

This time you can see the changes in the beginning of species distribution curves. Since the two excluded areas have high biological value, the proportions of species distributions decrease more steeply in the beginning when these areas are removed. However, since other valuable areas can be included in the solution, this masking has less effect on the final 15% top fraction compared to the old reserves, as you can see from the average and minimum proportions of species distributions.

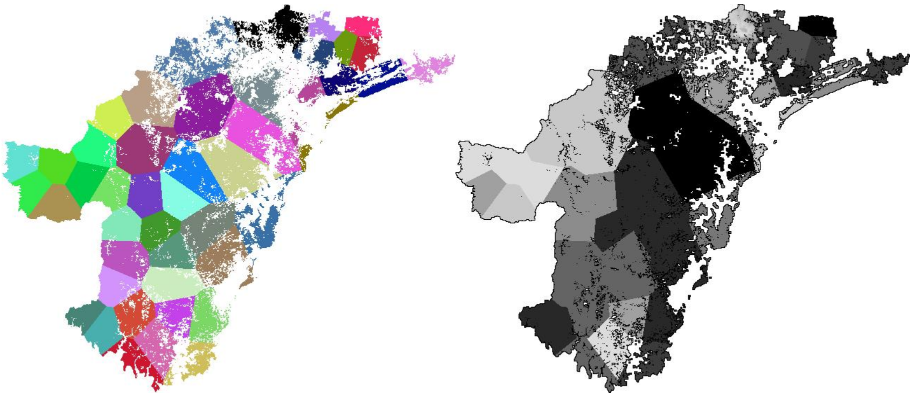


Average and minimum performances when residential areas are excluded from the top fraction.

6.8 Exercise 8

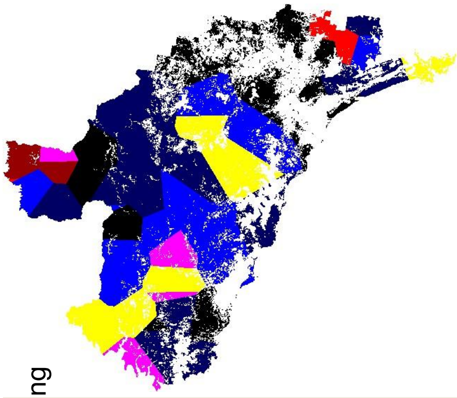
Working with directed connectivity

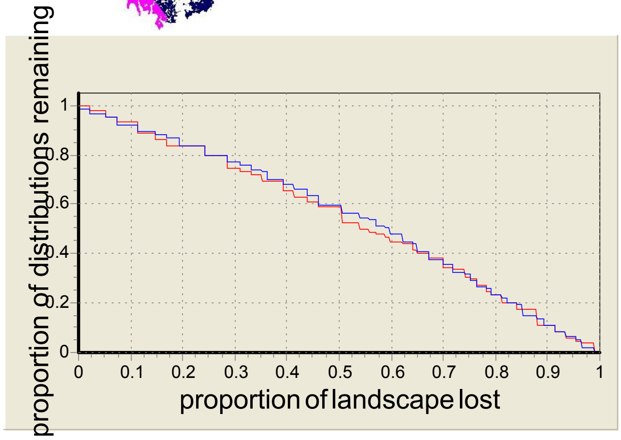
Next, we familiarize ourselves with the directed connectivity feature, NQP (Neighborhood Quality Penalty). This is a variant of BQP and thus the basic principles for landscape ranking are the same as in BQP. The essential difference is that here the connectivity of sites is clearly directed and defined as a set of linked planning units (instead of a buffer area around the focal cell as in basic BQP). To illustrate this, we now forget all the previous analyses we done so far and assume that the seven species we want to protect are in fact freshwater species that live in the numerous rivers in our study area. Thus, to account for connectivity while selecting sites for conservation actions, we need to consider both upstream and downstream connections of a site.



Picture of the study area divided into planning units (color) and river basins (grayscaled).

Because NQP works with planning units instead of grid cells, we have created a planning unit layer of the area called **plu_file.asc**. Planning units can be defined by any criteria, so lets assume that in our case they describe the smaller water catchment areas of riverine systems. The direction of water flow has been defined in a file **rivers.txt**, where every single planning unit has been linked to the following downstream unit. The penalty curves are given in the **NQPcurves.txt** -file, where we have a general curve for upstream and downstream connections, respectively. Settings for running the analysis are in **set_nqp.dat** and, in this case, we use the original species list file, **splist.spp**. To execute the analysis you can use the batch file **do_nqp.bat**.




proportion of distributions remaining
1
0.8
0.6
0.4
0.2
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
proportion of landscape lost

As you can see, the use of planning units changes the results quite radically. Also species performance curves change from smooth to staggered lines and in general all species are performing notably worse as compared to e.g. the basic analysis in Exercise 1. This is due to the fact that the planning units used here are relatively large and as the program removes units from the landscape, the risk of loosing valuable areas inside the units is very high. Therefore it is important to understand the use of large planning units will automatically cause a decrease in the quality of results and thus the sizes of planning units should be carefully selected.

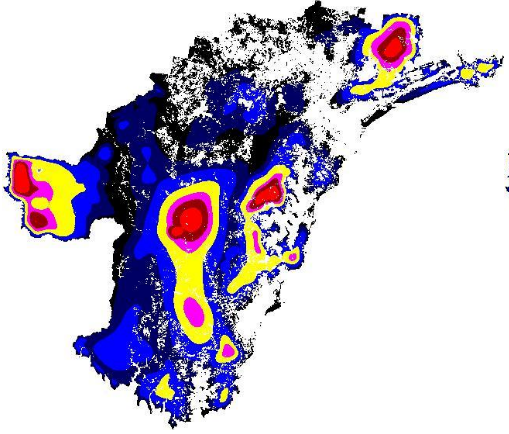
6.9 Exercise 9

Administrative units

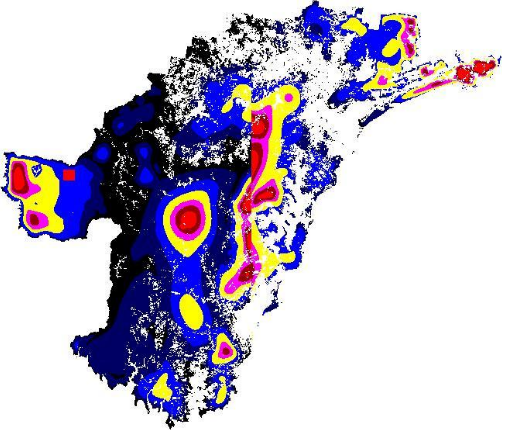
In this exercise our planning region comprises three different administrative units called Eastland, Westland and Boxland. The tiny Boxland is located within Westland and appears as a small red box in the strong variant images. The regions are defined in your ADMU layer file (ADMU_map.asc). For starters we assume all species are considered of equal value both globally and within each region (all weights equal in splist_ADMU.spp and ADMU_weights.txt). You may want to use connectivity to obtain same results as in the example images where (the third last command line parameter =1). We also start with considering the regions equally valuable (G_A is 0.33 for all regions). Select the appropriate settings in your set_ADMU.dat file to test the different variants. With the weak variant the solution does not differ from an analysis that does not distinguish between separate areas, whereas with the strong variant, especially when giving almost all value to local considerations (p=0.001), each region clearly has its own independent priorities. Increasing the *p* value will give increasing emphasis to the global component, and the result will again start to resemble a global solution.

Batch file do_ADMU.bat

All weights equal, weak variant

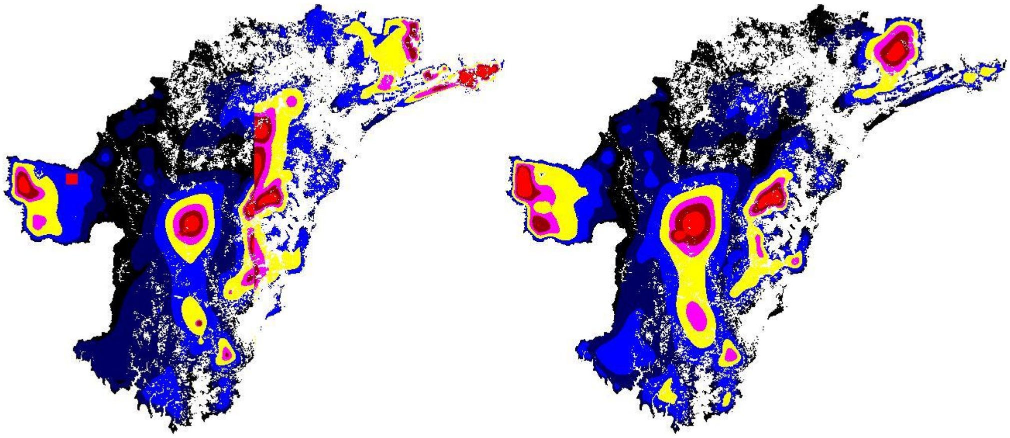


All weights equal, strong variant, p=0.001



All weigths equal, strong variant, p =0.8

Weak variant, $w^G_{Eastland}$ = 0.5, others 0.25



Let's imagine that Eastland is a region with a small human population, and therefore lower threats to its biodiversity, and has its habitats in better condition. Assuming that the species would be safer when protected there than in the other regions, we want to give a double weight to Eastland (in your ADMU_descriptions.txt file $G_{Eastland}$=0.5, $G_{Westland}$ and $G_{Boxland}$ are 0.25). (Note that one might equally well prefer to prioritize regions most under pressure and ignore the ones where the species will be safe anyway).

With strong priorities representation is required for all species within all regions regardless of their size, which may produce politically unacceptable outcomes, such as prioritizing the entire Boxland in our example. Area-standardized regional weights are one option to alleviate this problem. Try assigning area-standardized weights (G_A in the descriptions file) of 0.739, 0.259 and 0.002 to Westland, Eastland and Boxland, respectively. To see the effect best, run it with a low *p* value, such as 0.001.

Modify the global and regional species weights to see their influence. When species weights are not equal, the $ß_A$ parameter comes into use (so far its value has not had an influence because all species weights have been one). Try adjusting it as well. Modify one thing at a time to keep the changes in results tractable.

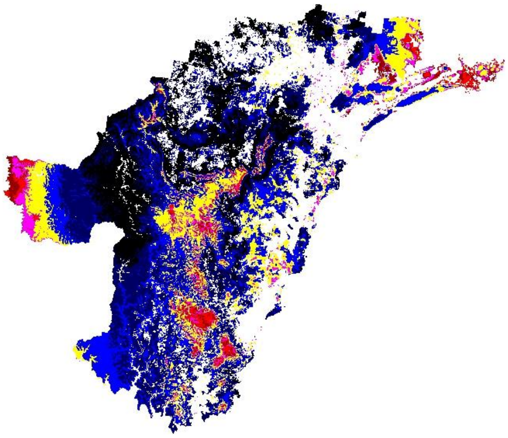
6.10 Exercise 10

Community level analysis with similarity expansion

In this exercise, we will take a community-based approach to reserve selection. We will use three layers representing different community types:
community1.asc
community2.asc
community3.asc

To be able to find the outputs easily, we assign them in a folder called **out_community**. Please create a folder with this name to your directory or adjust all the batch files so that you get the output files where you want them.

First, we run the analysis with the three layers only. We need a list of the community type layers (**bdlist_community.spp**), a run settings file (let's use the very basic settings in **set.dat** for this run) and a batch file (**do_comm_nosimilarity.bat**).



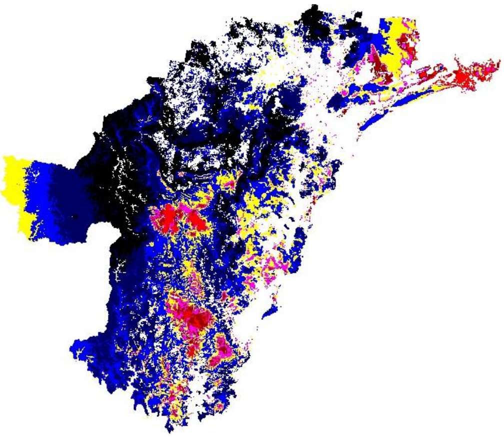
Curious to see how a community similarity matrix affects the result? Let's use another run settings file (**set_community_similarity.dat**). There we refer to a community similarity matrix called **community_similarity.txt** which is a basis for a similarity expansion (see section 2.8.). The matrix looks like this:

Community	1	2	3
1	1	0.6	0.1
2	0.6	1	0.1
3	0.1	0.1	1

Zonation calculates representation for community type 1 for one grid cell as follows:

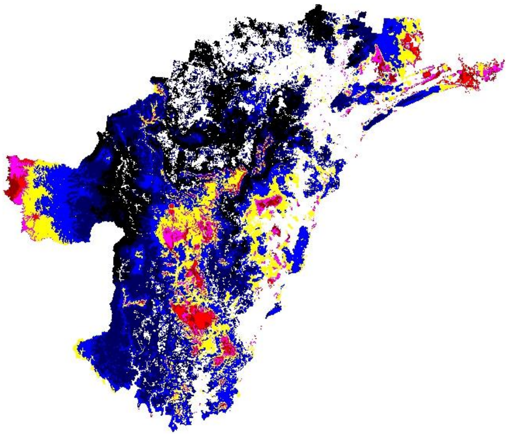
representation(community1) = occurrence(community1) + occurrrence(community2)*0.6 + occurrence(community3)*0.1

and treats the other community types in a similar way. We run the analysis with **do_comm_similarity.bat** and see how it affects the Zonation ranking of cells:



Please note that the representation levels in the **.curves.txt** output file are calculated for the expanded occurrences of community types (the values calculated with the similarity matrix). We can find out how the original community types are represented in the solution using solution loading (see section 3.5.2.). We call the program with **do_comm_original_rep.bat** where we ask Zonation to rank the cells exactly as it did in the previous analysis but calculate representation levels for the original community types. This analysis produces a file called **output_commsim_original.curves.txt** from which we can see how the representation of the community types themselves changes during cell removal.

We can also run an analysis where we take both community types and individual species into account. We make a biodiversity features list file (**sp_community_list.spp**) where we first list community types and then individual species we want to include in the analysis. We also want to have a grouped output to be able to check how representation changes during cell removal for community types and species separately. We assign community types in the output group 1 and species in output group 2 in a groups file (**groups_community_sp.txt**). We adjust the settings slightly and use **set_community_sp.dat** as our run settings file. We call Zonation with **do_comm_sp.bat** and look at the output:



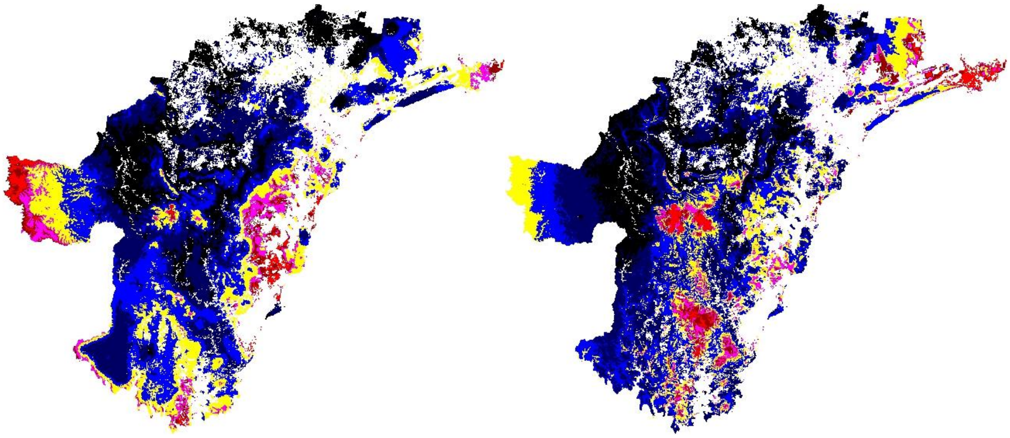
As we assigned community types and species into separate output groups, we now have a file called **output_comm_sp.grp_crvs.txt** among the outputs. This file summarises minimum, mean and maximum representation of both output groups separately. The representation curves of individual features are described in the **.curves.txt** file, as usual.

6.11 Exercise 11

Landscape condition and retention

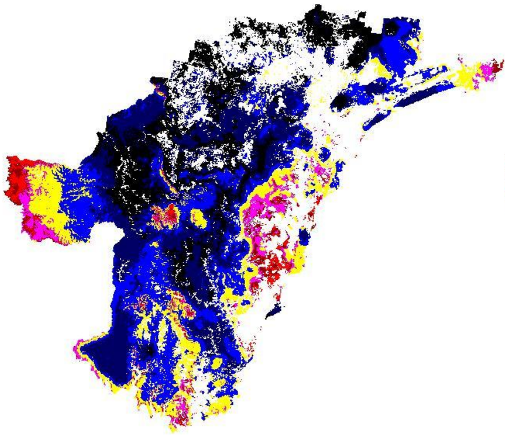
It is time to look at how landscape condition and retention affect the conservation priorities. For these analyses, we use the same imaginary community types as in exercise 10. Community types are often the most practical units for this analysis type. We also take community similarity into account. We have defined **out_cond_ret** as an output folder for all the analyses in this exercise, so please create a folder with that name in your directory or adjust the batch files so that the outputs will go where you want them to.

Let us first run an analysis with landscape condition. Remember from the previous exercise that community types 1 and 2 are rather similar while community type 3 is very different from both of the other two. We assign community types 1 and 2 into condition group 1 and community type 3 into condition group 2 in a groups file (**groups_condition.txt**). After that, we still need a condition layer list file saying that for condition group 1, landscape condition is described in **condition1.asc** and for condition group 2 in **condition2.asc**. In the run settings file (**set_condition.dat**) we state that we are using a groups file, a condition layer and a community similarity matrix. We call Zonation with **do_condition.bat** and see how condition affects the ranking:

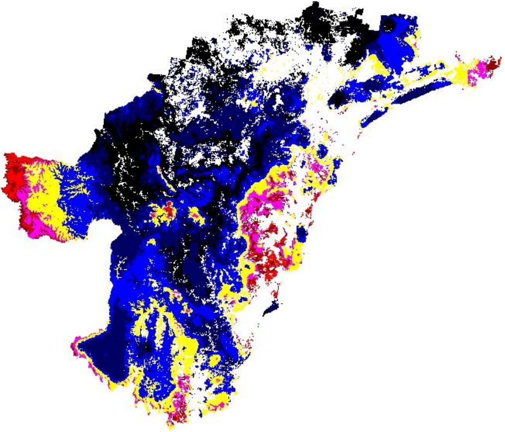


The figure on the left is the ranking with community similarity settings and with a condition layer. On the right, for comparison, is the ranking with community similarity but no condition. As community types 1 and 2 have good habitat condition (little past habitat loss) in the capes and in the south, their conservation value is increased in those areas. For community type 3, habitat condition is highest at the western edge, which increases the conservation value for that region.

Finally, let's apply retention analysis on top of what we have so far. Retention indicates the difference we can achieve with conservation action. The trick in retention analysis is to balance the representation of features in reserves and their retention across the full landscape. Therefore, we use a biodiversity features list file where we listed all community types twice (**bdlist_retention.spp**). For this analysis, we assume that all three community types have a similar pattern of retention, defined in **retention1.asc**. We list that layer in **ret_list.txt.** In **groups_retention_m1.txt** we state that i) we want to have separately grouped outputs for the layers used to calculate representation (the first set of layers) and those used to calculate retention (the second set of layers), ii) that we apply condition on both sets of layers, iii) that we apply retention only to the second set of layers and all of them belong to retention group 1 and iv) that we use retention mode 1, "stop loss", for all features. The settings are defined in **set_retention_m1.dat**. Finally, we use **do_retention_m1.bat** to call Zonation and run the analysis.



Condition and retention



Condition only

The figure on the left shows Zonation ranking when retention is accounted for, whereas the figure on the right shows the ranking with condition only, for comparison. See how the middle parts are gaining priority. This happens because retention is predicted to be high in the western areas, and therefore, the difference made with conservation action is higher in the middle and eastern parts.

Index

.

.curves.txt 129
.emf 129
.jpg 129
.prop.asc 129
.rank.asc 129
.run_info.txt 129
.wrscr.asc 129

A

across boundaries 232
adding aggregation 250
adding edge points 95
additive 31
additive benefit function 27, 29
administrative regions 73
administrative units 73, 232, 262
administrative units analysis 232, 262
administrative units analysis input 123
administrative units description file 123
administrative units files 123
administrative units map raster 123
administrative units setup 232
aggregation 250
aggregation methods 38, 39, 40, 41, 44, 172, 175, 178, 182
aim & purpose 10
algorithm 25
alpha value 90
alternative land use 67
alternative land uses 218
alternative land uses layer 121
alternative land uses setup 218
analyses 11
analysis area mask 117
annotations 95, 156
approximation, definition 13
areas included in the analysis 117
assumptions 76, 140
automated post-processing file 125

B

balancing alternative land uses 218
balancing representation and retention 221
base analysis 15
basic core-area Zonation 27
basic planning problems 194
basic Zonation 166, 242
batch file 17, 80
batch files 80
batch operation 80
benefit function 32, 90
biodiversity feature layer 121
biodiversity feature list file 90
biodiversity feature map files 87
biodiversity value 47
BLP 38, 39
BLP setup 171
boundary length penalty 38, 39

boundary length penalty setup 171
boundary quality penalty 38, 41, 175, 250
boundary quality penalty definitions file 106
boundary quality penalty setup 175
BQP 38, 41, 44, 106, 108, 175
BQP mode 95
BQP setup 175

C

call 80
cell removal principles 27
cell removal rule 27, 34, 90, 118
climate change 57, 228
climate change setup 228
colors 156
combined community and species level 213
command line 80
command prompt 17, 80
community and species 213
community level 64
community level analysis 209, 213
community level analysis setup 209
community similarity 110
community type distribution 87
community types 64, 209, 213
compensatory measures 207
competing land use 67
competition 57
compositional dissimilarities 64
compulsory input files 86
condition and retention 221
condition group 119, 121
condition layer 121
connectivity 38, 39, 40, 41, 44, 47, 57, 172, 175, 191, 207, 215, 250
connectivity across time steps 228
connectivity between time steps 72
connectivity in edges 111
connectivity in time 228
connectivity matrix 47
connectivity requirements 115
connectivity similarity 110
connectivity to multiple features 215
conservation planning 10
conservation prioritization 10
conservation value 49, 61
consumer-resource interaction 57
convex 31
convex optimization 12
coordinate files 104
core-area Zonation 27
cost 112, 138, 184, 255
cost curves 156
cost efficency analysis 255
cost efficiency 184
cost efficiency analysis setup 184
cost layer 95, 112
cost setup 184
curves 156
curves file 129

D

damaged areas 207

difference in representation 202

direct cost 67, 184

directed connectivity 38, 44, 95, 108, 118, 260

directed freshwater connectivity 178

directed freshwater connectivity setup 178

directory paths 143

dispersal alpha 90

distribution curves 156

distribution discounting 49, 50, 112

distribution maps 87

distribution smoothing 38, 40, 172, 250

distribution smoothing setup 172

distributional uncertainty map layer 112

download 17

downstream connectivity 44

dynamic landscapes 72, 225, 228

E

ecological interactions 57, 115, 191

ecological interactions setup 191

edge correction in connectivity 47

edge effect fix file 111

edge removal 95

errors in data 50

evaluating conservation areas 61

evaluating existing protected areas 202

evaluating proposed protected areas 202

examples 239

excluding areas 116

excluding areas from the solution 61

exclusion cost 61

exercise 1 242

exercise 2 245

exercise 3 248

exercise 4 250

exercise 5 253

exercise 6 255

exercise 7 257

exercise 8 260

exercise 9 262

existing protected areas 198, 257

expanding conservation areas 198

expanding protected area network 257

expected relative loss 69

extinction risk curve 156

extinction risk output 156

F

feature distribution 87

features 11

file names 143

file types 86

fragmentation uncertainty analysis 54

framework 11

G

GDM 64

generalized benefit function 27, 32

genetic algorithm 13
GIS 10
global weights 73, 232
graphical user interface 80
groups file 119

H

habitat loss 69
habitat loss rate 76
habitat model 10
habitat quality 69
habitat quality histogram 156
habitat restoration 225
habitat retention 69
habitat suitability 72
habitat types 64
heuristic 25
heuristic, definition 13
horizon of uncertainty 50
host-parasitoid interaction 57

I

identifying conservation priorities 194
incentive funding 205
including areas 116
including areas to the solution 61
inclusion cost 61
information gap theory 49, 50, 54
initial removal 95
input data 11
input file 104
input files 86, 104
installation 17
integer programming 13
interaction connectivity 72
interaction definitions file 115
interactions 95
interpretations 194
invasive species 57

L

land cost 112
landscape comparison 125
landscape condition 69, 121
landscape identification 125, 135
landscape identification for masked subregion 125
landscape identification for top fraction within masked areas 125
landscape retention 69, 122
landscape statistics 136
least conservation value 196
least harm for biodiversity 196
least valuable areas 196
limitations 13, 76, 140
literature cited 22
loading old solutions 85, 139
local administrative priorities 73
local conservation value 205
local edge correction group 119
local priorities 232
local weights 123
logit space 95
loss curve 41, 44
loss curves 106
loss function 41
lowest fraction 196
LSB 125

LSC 125
LSI 125
LSM 125

M

management intervention 69, 221
management landscapes 135, 136
management scenario 225
map colors 156
map layout 156
map window 156
maps -window 156
marginal loss 25, 27
mask layer 95, 116
matrix connectivity 47, 72, 182
matrix connectivity setup 182
memo window 156
memo -window 156
memory use 140
metapopulation connectivity 40
Metapopulation Research Group 17
mosaic of habitat types 111
MRG 17
multi-criterion analysis 67, 218
multi-feature ecological interactions 47
multiple costs 218
multiple features 47
multiple interacting features 215
multiple interactions 215
multiple time steps 72, 225, 228

N

name annotation 95
national 73, 232
national borders 47, 111
negative weights 67, 121, 218
neighborhood quality penalty 38, 44, 178, 260
neighborhood quality penalty setup 178
new features 19
new settings 95
nominal model 50
non-spatial Zonation 166
NQP 38, 44, 106, 108, 118, 178

O

objective function 12
offsetting 207
opportunity 49
opportunity analysis 54
opportunity cost 67
optimality 12, 13
optional input files 104
output 129
output files 129
output groups 119
overlap in landscapes 138

P

pair-wise similarities 64
Pareto-optimal boundary 50
past loss 69
planning units 38, 95, 118
point distribution data 104

pollution 57
positive surprises 54
positive uncertainty 54
post-processing analyses 135
power function 29
predator-prey interaction 57
printing options 156
proportion of habitat 47
proportional error 95
proportional selection 156
protected areas 207

Q

quality control 37
quick start 17

R

radius 41
random removal 37
rank file 129
ranking map 156
rare species 104
rarity map 156
raster map 87
references 22, 29
regional 73, 232
remaining map 156
removal mask 257
removal mask layer 198, 202, 207, 257
replacement cost analysis 61, 116, 202
representation 202
retention 221
retention analysis setup 221
retention group 119, 122
retention layer 122
retention mode 119, 122
reverse heuristic 25
richness 129
richness map 156
robustness 49, 50
run info file 129
run settings 95
running the program 80

S

scale of landscape use 90
selecting conservation areas 194
sensitivity analysis 15
setup combinations 236
setup for basic Zonation 166
setup for BLP Zonation 171
setups 194
similarity expansion 110, 209, 213
similarity matrix 64, 110, 209, 213
simple Zonation 166
simulated annealing 13
single-species prioritization 215
solution comparison 138
solution cross-comparison 85, 135, 139
solution cross-load 135, 139
solution loading 85, 139
species distribution 87
species info -window 156
species interactions 57, 95, 115, 191
species level analysis 213
species list file 90
species of special interest 104, 248

species-area curve 29, 156
species-cell-specific errors 112
species-info window 156
SSI species 95
SSI species list file 104
statistical distribution modelling 10
step function 31
stochastic optimization 13
stop loss 69, 221
subset of landscape 117
system requirements 140

T

target-based planning 27, 31
targeting of funding 205
targeting of incentive funding 205
testing old solutions with new settings 139
top fraction 194, 198, 207
top fraction outside existing protected areas 207
top fraction outside masked in areas 198
tree connectivity file 108
tree hierarchy 108
troubleshooting 143
tutorial 239
tutorial example 242, 245, 248, 250, 253, 255, 257, 260, 262

U

uncertainty analysis 49, 50, 54, 112, 186, 253
uncertainty analysis setup 186
uncertainty in distributions 50, 186, 253
uncertainty model 50
uncertainty parameter 50
uniform error 95
upstream connectivity 44

V

variable habitat suitability 225, 228
vector data 76
version 3 11
very rare species 248
visual output 156

W

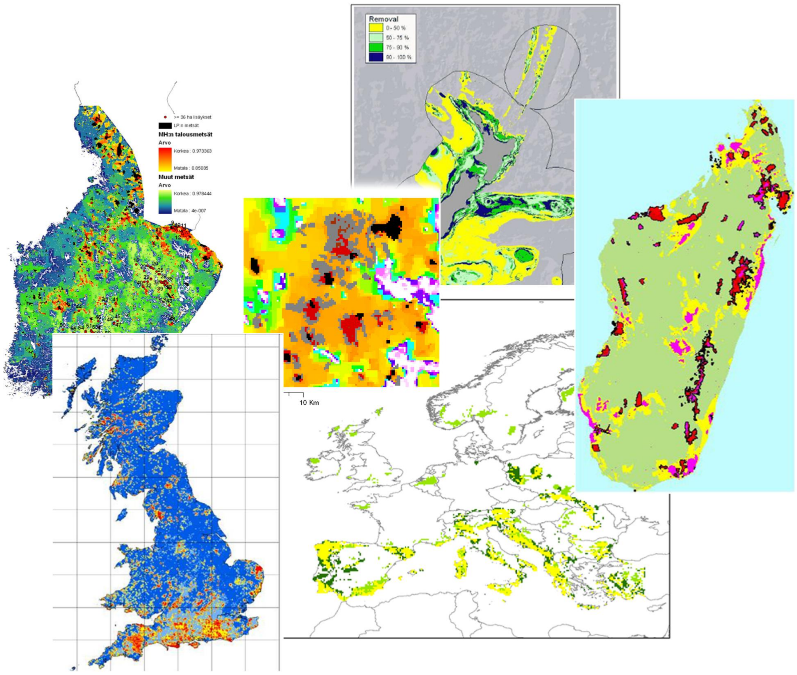
warp factor 95
water bodies 47
weighted range size corrected richness 129, 205
weighting of biodiversity features 245
weighting of species 245
weights 90, 245
weights matrix file 123
windows interface 17
wrscr 205
www 17

Z

Zonation compared to other reserve selection approaches 12
Zonation features 11
Zonation meta-algorithm 25

Zonation v.3.0 19, 22
Zonation workflow 15
z-value 95

For the latest version of Zonation, please visit our website
http://www.helsinki.fi/bioscience/consplan


Removal
0 - 50 %
50 - 75 %
75 - 90 %
90 - 100 %
>= 36 ha lisäykset
LP:n metsät
MH:n talousmetsät
Arvo
Korkea : 0.973363
Matala : 0.85085
Muut metsät
Arvo
Korkea : 0.978444
Matala : 4e-007
10 Km

The following partners have contributed significantly to the development and implementation of Zonation

UNIVERSITY OF HELSINKI
ACADEMY OF FINLAND
SCALES
European Research Council
erc

Metapopulation Research Group

Contact information:
Metapopulation Research Group
Biodiversity Conservation
Informatics Group

Department of Biosciences
P.O. Box 65 (Viikinkaari 1)
FI 00014 University of Helsinki
FINLAND
atte.moilanen@helsinki.fi